LIBRARY

APR 3 1928

Vol. XI, No. 2 *PSYCHOLOGICAL REVIEW PUBLICATIONS* APRIL, 1928

Journal of Experimental Psychology

EDITED BY

MADISON BENTLEY, UNIVERSITY OF ILLINOIS

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, Business Editor

CONTENTS

Remote Associative Tendencies in Serial Learning: MARGARET E. HALL 65
Comparison of Tonal Distances: CARROLL C. PRATT 77
The Concept of the Threshold and Heymans' Law of Inhibition: LLEWELLYN T. SPENCER 88
The Course of the Auditory Threshold in the Presence of a Tonal Background: ERNEST GLEN WEVER AND STANLEY R. TRUMAN 98
The Relation of Physical Constitution to General Intelligence, Social Intelligence and Emotional Instability: H. E. GARRETT AND W. N. KELLOGG 113
Galvanic Phenomena of the Skin: LLOYD A. JEFFRESS 130
A Study of the Learning Curves for Two Systems of Shorthand: PARRISH LITTLE JETTE 145
Oral Perceptions in Relation to Anosmia: H. R. CROSLAND, RUTH C. MILLER AND WINIFRED E. BRADWAY 161

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

Psychological Review Publications

of the American Psychological Association

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*J. of Exp. Psych.*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, Business Editor

WITH THE CO-OPERATION OF
MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in May, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

ANNUAL SUBSCRIPTION RATES

Review: $5.00 (Foreign, $5.25). Review and Bulletin: $9.50 (Foreign, $10.00)
Journal: $5.00 (Foreign, $5.25). Review and Journal: $9.00 (Foreign, $9.50)
Bulletin: $5.50 (Foreign, $5.75). Journal and Bulletin: $9.50 (Foreign, $10.00)
Any one of above with Index: $1.50 additional.
Review, Bulletin, and Journal: $14.00 (Foreign, $14.75).
Review, Bulletin, Journal, and Index: $15.00 (Foreign, $15.75).
Current Numbers: Review or Journal, $1.00; Bulletin, 60c; Index, $2.00
Psychological Monographs: $6.00 per volume (Foreign, $6.30).
Current Issues: prices vary according to size.

Subscriptions, orders, and business communications may be sent direct to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J. OFFICE

Journal of
Experimental Psychology

VOL. XI., No. 2 April, 1928

REMOTE ASSOCIATIVE TENDENCIES IN SERIAL LEARNING

BY MARGARET E. HALL[1]

Psychological Laboratory, University of Wisconsin

INTRODUCTION

Ebbinghaus put forward his theory of remote association in 1885.[2] In substance the theory maintains that, when a series of items is learned in a certain serial order, associations are formed not only between adjacent items but also, and at the same time, between items remote in the series. The theory was supported by an experimental investigation which was epoch-making because of its originality. One of the chief innovations introduced by Ebbinghaus was the use of nonsense syllables, which he used exclusively in his investigation. From the point of view of the present study it is particularly to be observed that Ebbinghaus learned his series at a single sitting and to one perfect reproduction. The series were then tested for remote associative tendencies after a lapse of 24 hours.

Recently the matter of remote association has been reëxamined both theoretically and experimentally by Cason.[3] This writer had his subjects learn a passage of prose by any method which they individually chose, and then recite it

[1] The writer wishes to express her indebtedness to Professor Clark L. Hull for suggestions and criticism during the course of the investigation.

[2] H. Ebbinghaus, *Memory; a contribution to experimental psychology* (trans. by Ruger and Bussenius).

[3] H. Cason, Specific serial learning; a study of backward association, this JOURNAL, 1926, 195–227; Specific serial learning; a study of remote forward association, *ibid.*, 1926, 299–324.

daily for a week. The passage was subsequently broken up into lists of words in remote order which the subjects relearned at the conclusion of the week of daily recitation. Cason compared the results of learning in the remote forward order with that of the control lists both before and after memorizing the prose passage. He found practically no evidence of remote association.

There are several factors present in Cason's experiment which may account for the failure to find remote association, and which, in any case, unnecessarily complicate the situation.

In the first place, logical material was used. This makes it impossible to tell how many and what previous associations have been formed with each word and in what directions these associations lead. In addition Cason (like Ebbinghaus) had his subjects learn the material from cards on which the lists were typewritten. The subjects saw all the words of a given list at one time, and although they were directed to read straight through the lists, their eye-movements could hardly have been sufficiently controlled to prevent a certain amount of skipping.

A second matter is that both Cason and Ebbinghaus in arranging their remote-order lists also removed the syllables from their original absolute positions in the series. While it would be difficult to say what effect this would have on saving, in learning the remote series, it is evidently an additional and unnecessary complicating factor. It would seem to be a sound principle of scientific method to avoid in experimentation, so far as is possible, all variable factors not strictly germane to the problem.

Moreover, it is evident that in Cason's experiment a large amount of overlearning took place because the subjects, after learning the passage, recited it daily for a week. This overlearning, added to the immense amount of learning from years of use of the language, may have interfered with the functioning of remote associative tendencies when the crucial test was made.

It is also to be observed that, even when material is barely learned, interference may still take place if the relearning of the remote-order lists is done immediately or soon after the original learning. Ebbinghaus in his investigation allowed 24 hours to elapse between the first learning and the relearning. This may be one reason why he found evidence of remote association. He himself seemed, however, to have been unaware of any special influence of the forgetting period. Neither does Cason seem to recognize either overlearning or lapse of time before relearning as possible factors determining the manifestation of remote associative tendencies. The present writer wishes to suggest the possibility that after there has been a period of forgetting, interference may disappear because of the weakening of bonds between consecutive items, and that remote association may then manifest itself. Moderate original learning together with a period of forgetting may be necessary for the functioning of remote associative tendencies.

Because of the various considerations mentioned above, it was decided in the present investigation to test for remote association both immediately and after a week of forgetting

had taken place. Accordingly, each subject relearned half of the test material and half of the control material *immediately*, and the remaining half of each kind of material *after one week*. In this way it was possible to compare the results of immediate relearning with those from the delayed relearning upon remote associative tendencies.

The main problem of this investigation, then, is to see whether remote associative tendencies exist at all; and, if so, whether or not they tend to appear more strongly at once or after a period of forgetting.

The Experiments

The subjects were 40 university women students. Very few of them had studied psychology at all, and none of them were familiar with the alleged law under investigation.

Hull's automatic exposure apparatus was used. By raising a small window in the face of the device one syllable can be exposed at a time. A drum revolving behind carries a canvas band on which three lists of syllables are printed in bold gummed letters ¼ inch high. The syllables of a given list appear one after another at the window, each for exactly 1.5 sec. After the last syllable has disappeared, a blank space appears for 3 sec. and then the list starts over again, and the drum continues to rotate until the subject has learned the series to the point of one complete correct recitation. The experimenter then closes that window, opens the adjoining one, and the same process continues for the next list, and so on.

An exposure apparatus was used in order to prevent eye movements back and forth along the list, which were impossible to prevent with the method employed by Cason and Ebbinghaus. Moreover, each syllable was exposed only 1.5 seconds. While this does not entirely eliminate "thinking" back and forth, it reduces it considerably, because with this rate of the exposure the subject is kept very busy merely responding.

Material.—In the present investigation nonsense syllables were used as the learning material. Although a few previous associations probably complicate the behavior based on such material, the number of associative tendencies thus engendered is much smaller than from logical material, and the learning is put on a more stable and uniform basis for all subjects. Moreover, nonsense material can be more easily controlled and manipulated for experimental purposes than can logical material. And again, learning begins more nearly at zero level.

There were used in all 120 nonsense syllables of three letters each. Care was taken to eliminate any which appeared to have meaningful associations. They were

arranged in 12 lists of 10 syllables each. These 12 lists will be referred to as the *original forward* lists. They are referred to individually as A, B, C, D, E, F, G, H, I, J, K, and L. The 12 lists were divided into 4 groups of 3 lists each, the groups being indicated as I, II, III, and IV.

In making up the remote test-lists alternate items were used. It was assumed that if remote association exists at all, it ought to be strongest between alternate items. After the negative results of Cason, it was thought best to seek evidence at the most favorable point.

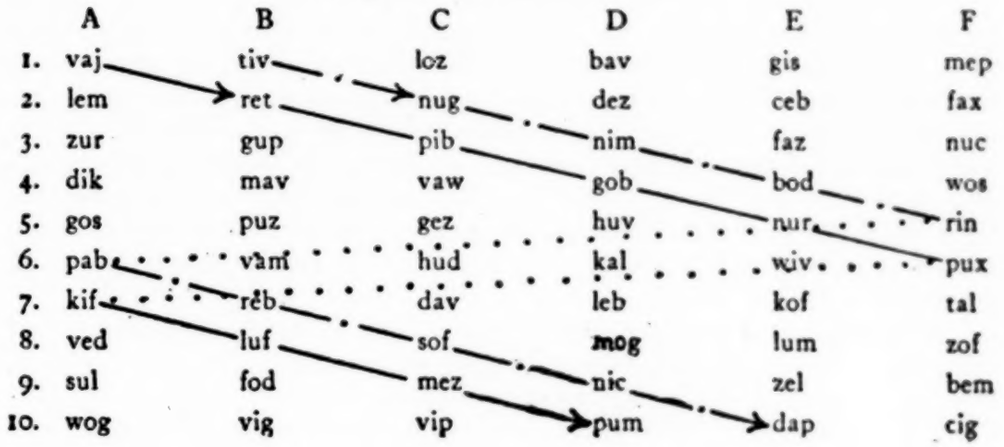
The method of making up the remote test-lists will be explained by using Lists A, B, and C of Group I by way of illustration. The remaining groups of syllable lists were treated in the same way. In making up the remote test-lists, the even-numbered items of A were omitted and the even-numbered items of B substituted in their places. This left gaps in B, and in these gaps were substituted the even-numbered items of C. The gaps thus left in the C list were then filled in by the even-numbered items of the A list. The details may be seen in the typical series as follows:

	Original Lists I			Remote Test-Lists I′		
	A	B	C	A′	B′	C′
1.	vaj	tiv	loz	vaj	tiv	loz
2.	lem	ret	nug	ret	nug	lem
3.	zur	gup	pib	zur	gup	pib
4.	dik	mav	vaw	mav	vaw	dik
5.	gos	puz	gez	gos	puz	gez
6.	pab	vam	hud	vam	hud	pab
7.	kif	reb	dav	kif	reb	dav
8.	ved	luf	sof	luf	sof	ved
9.	sul	fod	mez	sul	fod	mez
10.	wog	vig	vip	vig	vip	wog

Thus all the syllables in the 3 original lists were utilized in the remote test-lists without changing the position of a single individual syllable in any of the series. By this arrangement each syllable has the same position in the series of 10 that it had in the original forward-list; but is arranged so that all immediate bonds of the old list are omitted. These new lists will be referred to as *remote test*-lists.

In making up the control lists it was, of course, necessary to put syllables together which had not been associated in the original lists. It was also deemed desirable, as mentioned above, to have each syllable keep its same relative position in the series of 10. The method by which this was accomplished will be illustrated by using 6 original lists A to F. The control-lists were formed by taking the syllables diagonally down from left to right across the six lists. For example, control list A* was made up by taking the 1st syllable of A, the 2d of B, the 3d of C, the 4th of D, the 5th of E, the 6th of F, the 7th of A, the 8th of B, the 9th of C, and the 10th of D. It will be observed that it was found necessary to double back for the last 4 syllables, but there is a gap of 5 syllables between the 1st taken from A and the 7th taken from A. It is to be noted in addition that any remote association which might exist between the 1st and 7th would hinder rather than facilitate the proof of the existence of remote association by decreasing the difference between the results of learning the test- and the control-lists.

Original Forward Lists



	A	B	C	D	E	F
1.	vaj	tiv	loz	bav	gis	mep
2.	lem	ret	nug	dez	ceb	fax
3.	zur	gup	pib	nim	faz	nuc
4.	dik	mav	vaw	gob	bod	wos
5.	gos	puz	gez	huv	nur	rin
6.	pab	vam	hud	kal	wiv	pux
7.	kif	reb	dav	leb	kof	tal
8.	ved	luf	sof	mog	lum	zof
9.	sul	fod	mez	nic	zel	bem
10.	wog	vig	vip	pum	dap	cig

Control Lists

	A*	B*	C*	D*	E*	F*
1.	vaj	tiv	loz	bav	gis	mep
2.	ret	nug	dez	ceb	fax	lem
3.	pib	nim	faz	nuc	zur	gup
4.	gob	bod	wos	dik	mav	vaw
5.	nur	rin	gos	puz	gez	huv
6.	pux	pab	vam	hud	kal	wiv
7.	kif	reb	dav	leb	kof	tal
8.	luf	sof	mog	lum	zof	ved
9.	mez	nic	zel	bem	sul	fod
10.	pum	dap	cig	wog	vig	vip

As a second example list B* is made up by taking the 1st syllable of B, the 2d of C, the 3d of D, the 4th of E, the 5th of F, the 6th of A, the 7th of B, the 8th of C, the 9th of D, and the 10th of E. The other control lists were formed in a similar manner.

We now have three kinds of lists: original, test, and control, and twelve of each kind.

Procedure.—1. *Learning Conditions.* The Ss all did their learning during daylight in the same room and sitting in the same position before a table on which the apparatus was placed. The writer always acted as experimenter, and sat at the side of the table checking the responses. The four learning periods of each subject were kept constant as to time of day.

2. *Procedure in general.* When the subject entered the room she was told to sit down in the chair and was handed the following printed directions to read.

"This is a memory test. You are to learn these nonsense syllables in the right order as quickly as you can. I will open the first window. When I start the machine the syllables will appear in it one after another until ten have been shown. Try to learn them as they appear. After the tenth syllable a blank space will appear. Then the same list will start over again. While the blank space is there try to name the first syllable, (not spell it), before it appears, and the 2d one just before it appears, etc. Name each syllable while the one before it is being shown. Guess at it even if you are

not sure, but be careful not to name it after it actually appears, because that would be reading, not memory. The list will be shown over and over again until you have learned it. Be sure that you are not making some slight mistake in a syllable you think you know. Pronounce the syllables just as they are spelled and try to be consistent each time. Read the directions again if you do not understand them." The experimenter started the apparatus and checked as + or − the responses of S. When S had once recited the list correctly, the experimenter closed the window and immediately opened the next. S was told to learn this second list just as she had learned the first. By this method S automatically checks up on her own errors, and no verbal prompting is required from the experimenter. It is at once a checking up and a recitation.

3. *Specific procedure for one Subject.* The first day S learned 1 group of 3 original lists, rested 5 minutes, learned another group of 3 original lists, again rested 5 minutes, and then relearned the first group in *remote* test-form.

The second day at the same hour S learned a third group of 3 original lists, rested 5 minutes, learned another (the fourth) group of 3 original lists, again rested 5 minutes, and then relearned the third group in *control* form. Thus, there is a perfect balance as to procedure, amount of material, time order, etc., between the 1st and 2d day's learning, except that the 1st day the relearning was done with test-material and the 2d day with control-material.

The eighth day and at the same hour S relearned the second group (originally learned the first day) in remote test-form.

The ninth day at the same hour S relearned the fourth group (originally learned the second day) in *control* form. Here again there is a perfect balance between the two latter relearnings. Each was done just one week after the original learning, and in both cases the amount of material and the procedure were the same. This subject's work is shown in schematic form in Program 1 of Table I.

Table I

The Eight Learning Programs

Roman numerals represent groups of 3 lists each. Without any marking they represent the original *forward* groups. With a prime mark they represent groups of lists arranged in *remote test* order. With an asterick they represent groups of lists arranged in *control* order. Each program was carried through by five subjects.

Learning Periods	Programs							
	1	2	3	4	5	6	7	8
1st Day	I II I′	I II I*	III IV III′	III IV III*	II I II′	II I II*	IV III IV′	IV III IV*
2d Day	III IV III*	III IV III′	I II I*	I II I′	IV III IV*	IV III IV′	II I II*	II I II′
8th Day	II′	II*	IV′	IV*	I′	I*	III′	III*
9th Day	IV*	IV′	II*	II′	III*	III′	I*	I′

The procedure just described was typical of that used for all Ss. The material was rotated, however, so that half of the Ss relearned I and II as test-material and III and IV as control-material, and half relearned I and II as control-material and III and IV as test-material. Half the Ss learned test-material the 1st day and control-material

the 2d day, and half reversed this order. Eight different learning programs were made out to take care of the various possible constant errors and 5 Ss learned each program. It will be seen from these programs (Table I) that, in the final results of the 40 Ss, neither the test- nor the control-material received any advantage from ease of material, place in the learning program, procedure or practice effects, nor any disadvantage from difficulty of material or fatigue. Any possible advantage which one program might give to one kind of material would be cancelled out by the use of the 8 balanced programs.

Results

The number of repetitions required to achieve one correct recitation was taken as the measure of efficiency of learning. The results of the investigation are summarized in Table II, where it will be seen that in the immediate relearning there is a saving in repetitions of 32.3 per cent. for the test-lists and 30.7 per cent. for the control-lists over the original learning of these lists, showing a difference of 1.6 points in favor of remote association. The comparison was put on a percentage basis to equalize any differences in the original learning.

Table II

The Final Results on Remote Associative Tendencies

	Original Learning				Immediate Relearning		Relearning After 1 Week	
	Test		Control		Test	Control	Test	Control
Sum	1856	1566	1765	1586	1257	1222	1141	1263
Mean	46.4	39.15	44.125	39.65	31.425	30.55	28.525	31.575
S.D. of distrib'n					10.936	13.762	8.193	9.767
% Saving over orig. learning					.322	.307	.271	.203
Diff'ce between means					.875		3.05	
$P.E._D$					1.5586		.6515	
Correlation coefficient					.7217		.7898	
$\frac{D}{P.E._D}$					.5614		4.6815	

The final result is that, in the *immediate* relearning, the test-lists show 4.9 per cent. more saving over the original-lists than do the control-lists. This was found by dividing the difference between .322 and .307 by .307. In the relearning *after a week* the test-lists show 33.4 per cent. more saving over the

original-lists than do the control-lists. This was found in a similar manner by dividing the difference between .271 and .203 by .203. It means that the test-lists were only slightly better than the control-lists in the immediate relearning, and very much better in the relearning after a week.

In more general terms, remote associative tendencies appear only very slightly when tested by immediate relearning. If the investigation had stopped at this point, as did Cason's, there would be no very conclusive evidence in favor of remote association. This study was carried farther, however, and the results from relearning after a week seem to point toward the conclusion that remote associative tendencies are rather strongly effective after a period of forgetting.

Because the columns of data compared were obviously correlated, the following formula using the correlation coefficient was used for the probable error of the differences.[4]

$$\text{P.E.}_D = .67449\sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}{n}}.$$

To find the reliability of the differences, the difference was divided by the P.E._D. The reliability of the difference between the test- and control-lists in the *immediate relearning* is .5614. The smallness of this number does not indicate that the difference is not reliable. It merely fails to show either one way or the other whether or not there is a real difference.

The reliability of the difference between the test- and control-lists in the *relearning after a week* is 4.6815, which shows a highly reliable difference. Theoretically, this indicates that there is only about one chance in a thousand that such a difference could have been produced by chance.

Conclusions

The results of the present investigation suggest that there are two external factors to be considered in a discussion of the evidence for remote association. They are (1) the factor of *degree* of learning and (2) the factor of *lapse of time*, a period of forgetting. In reality these two factors are probably both aspects of the same physiological process. It is in respect to these two considerations that the studies of Cason and Ebbinghaus greatly differ. Neither of these writers seems, however, to have recognized the possible effects of these

[4] W. H. Winch, Inductive versus deductive methods of teaching; an experimental research, p. 9. (*Educ. Psychol. Monog.*)

factors upon remote associative tendencies. Their results, then, in the light of the above considerations, may not really be in conflict, as Cason has supposed.

A fundamental principle generally accepted by psychologists assumes that after a stimulus-pattern has been conditioned to a response-pattern, part of the stimulus-pattern may then evoke the whole response-pattern. The original lists of 10 syllables in this study are probably stimulus-response sequences of the general nature suggested above. In both the test-lists and the control-lists, each syllable tends to evoke the original sequence to which that syllable had been conditioned. The item which followed it in the original sequence tends to be evoked. But when, instead, an item not previously conditioned to it now follows, one would naturally expect a conflict which would make the formation of a new bond more difficult. The old conditioning has to be broken down before a new one can be formed. In this respect both the test- and the control-lists are alike; a new bond has to be formed after each syllable and an old bond broken.

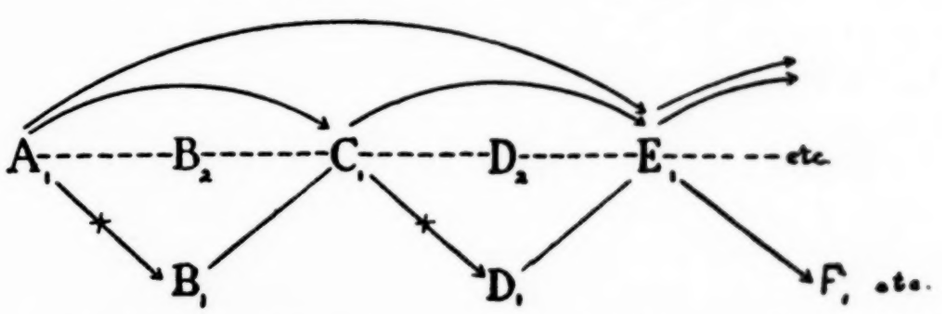


FIG. 1. The series of letters represents a *test*-list, 1's representing the syllables taken from one original list, 2's representing syllables taken from another original lists In the following discussion, only the sequence represented by 1's will be taken into consideration. It will be understood that the other sequence, represented by 2's, act. in the same way. Solid lines represent associative tendencies already formed. Dotted lines represent 'bonds' which are to be formed, and lines with a cross represent 'bonds' which must be broken. The letters below the horizontal line represent old syllables formerly in that position but no longer present.

Except in this particular, the two kinds of lists are different. In a control-list each syllable is taken from a different original sequence. In a test-list, on the other hand, all the syllables are taken from either one or the other of two original se-

quences. In other words, every alternate item in a test-list comes from the same original stimulus-response sequence.

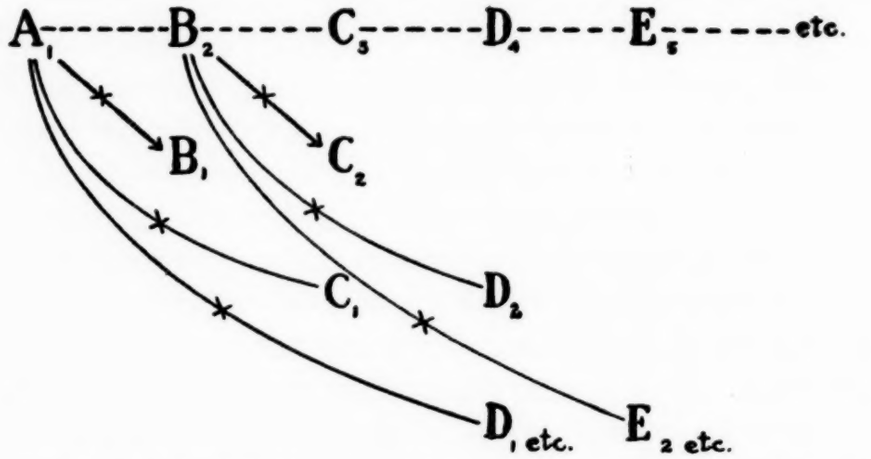


FIG. 2. This series of letters represents a control-list, each numeral indicating the original, forward list from which the syllable was taken.

The immediate relearning will be discussed first with the aid of the diagrams shown in Figs. 1 and 2. In the presentation of the first two syllables the conditions would be the same for both test- and control-lists. The syllable A_1 would tend to evoke B_1, C_1, D_1, etc., and B_2 would tend to evoke C_2, D_2, E_2, etc. In a control-list none of these syllables which tend to follow are actually presented later as a stimulus. The 'bonds' A_1 B_1 and B_2 C_2 must be broken, and the new 'bond' A_1 B_2 must be formed. In a test list (Fig. 1) the tendency of A_1 to evoke B_1, C_1, D_1, E_1, etc., facilitates the relearning of C_1 when C_1 is actually presented later, because there has, according to our experimental results, been a 'bond' previously established between A_1 and C_1 as well as between A_1 and B_1. When C_1 actually appears it brings a tendency to evoke D_1, E_1, F_1, etc. This tendency is stronger than the tendency of A_1 to evoke B_1, C_1, D_1, etc., because C_1 is presumably reinforced by A_1, which has occurred just previously. C_1 and A_1 together form a larger part of the old-stimulus-response sequence than A_1 or C_1 alone, as would be the case in the control-lists. It is assumed here that the more completely an original stimulus pattern recurs, the stronger will be the tendency to evoke the whole response-pattern or sequence.

There is at C_1, then, a stronger tendency to evoke the syllables which originally followed than there would be at the 3d syllable (C_3) of a control list. In a test-list the associative tendency C_1 E_1 can be utilized, because E_1 is in the test-list; but the tendency C_1 D_1 has to be broken. Because some of the old syllables which tend to be evoked in the test-lists are actually used farther on in the same lists, the learning of the test-lists is greatly facilitated. This is evidently not the case in the control-lists. At the same time, the associative tendencies between items formerly *immediate* have to be broken in the test-lists, and this means a strong inhibitory or interfering tendency in the learning of these lists. In the control-lists there is also an inhibitory tendency; but it is weaker because each syllable comes from a different original sequence, and, therefore, stands alone. There is not the reinforcement which is found in a test-list.

To summarize, this would mean that there are these two tendencies (inhibitory and facilitating) present in both the test- and the control-lists. They both seem to be present only slightly in the control-lists, and both present rather strongly in the test-lists, as explained above. In the *immediate* relearning the *ratio* between the two processes seems to be about the same, however, for both kinds of lists, because the results show the efficiency of learning both kinds of lists to be about the same.

During the week of forgetting, the 'bonds' of the original stimulus-response sequences have weakened considerably, and the results of the present investigation seem to show that the immediate bonds tend to weaken in greater proportion than the more remote bonds. This would mean, then, that the inhibitory tendency must weaken in greater proportion than the facilitating tendency. After a week of forgetting, the *ratio* between these two tendencies has evidently changed from what it was at the time of the immediate relearning. In the test-lists the lapse of time may have weakened the immediate inhibitory tendencies faster than the remote facilitating tendencies, and remote associative tendencies may have had a chance to function thus producing the saving. In

the control-lists also the *ratio* probably changed but the facilitating tendency was originally weaker (as pointed out above) than that found in the test-lists. This may account for the greater saving of the test-lists after a week of forgetting.

If the material had been overlearned, as in Cason's study, the difficulty of breaking up the habit sequences and recombining them in new ways would have been greater than in this study where the material was but barely learned. In that case it would presumably have taken longer for the inhibitory tendencies to fade out, and it is probable that remote associative tendencies would never have been so effective as with the barely learned material.

The final result, then, is that the test-lists are very little better than the control-list in the *immediate* relearning and very much better in the *delayed* relearning. This would lead to the ultimate conclusion that *the functioning of remote associative tendencies is inhibited in proportion to the strength of the original learning and is facilitated, within certain limits, in proportion to the length of the period of forgetting.* It is easier to break up and recombine parts of a *weakly* established than parts of a *strongly* established stimulus-response sequence. The weaker the original sequence, the more easily it is broken up and recombined in new ways. Moreover, the weaker the original sequence becomes through lapse of time, the more easily will its parts become parts of new stimulus-response sequences.

(Manuscript received Sept. 25, 1927)

COMPARISON OF TONAL DISTANCES

BY CARROLL C. PRATT

Harvard University

Prior to 1875 it was believed, naturally enough, that the same musical interval represents a constant relative pitch-difference no matter from what part of the scale it may be taken. The publication of Preyer's work [1] on the j.n.d. of pitch tended, however, to undermine this conviction in certain quarters.

Thus Langer wrote: "Immerhin ist die Annahme, dass in dem Hören von Octaven gleiche Empfindungsunterschiede abgeschnitten werden, eine willkürliche Annahme, welche nur ein Ausdruck für eine Periodicität ist, die in gewissen Relationen der Schwingungszahlen begründet ist." [2] Müller was also doubtful about the correctness of the general belief, "denn erstens möchten wir nach unseren Erfahrungen thatsächlich nicht einmal die Behauptung wagen, dass ein Intervall, dem ein constantes Verhältniss der Schwingungszahlen entspreche, auf verschiedenen Höhenstufen wirklich immer gleich gross erscheine." [3] Finally Stumpf, after an enumeration of the difficulties involved in making comparison of intervals with respect to their size, asserted that according to his judgment "das gleiche Intervall mit zunehmender Tonhöhe bis etwa c^3 eine zunehmende Distanz in der Empfindung darstellt." [4]

This divergence of opinion from the more natural assumption was based on very little factual evidence. There were no experiments, in fact, dealing directly with the problem of interval-comparison. Doubt as to the psychological equality of the same interval from different parts of the scale seemed to derive rather from the suspicion that there might be a slip somewhere between the apparent *musical* equality of intervals with the same vibration-ratios and the actual *psychological* equality of these same intervals judged in terms of the amount of difference between their limiting components To the musician it seems obvious enough that all octaves, or fifths, or thirds, *etc.*, are the same whether they are from high treble or low bass. But is it equally obvious, to the psychologist, that the amount of difference in auditory quality between the components of the interval 32–40 and the interval 1024–1280 (both major thirds) is the same? Not until the publication of Lorenz' paper [5] in 1891 was there any direct experimental evidence on this question. It is not the place here to summarize Lorenz' results or to consider the heated arguments between Stumpf and Wundt which arose over the

[1] W. Preyer, *Die Unterschiedsempfindlichkeit für Tonhöhen*, 1876.

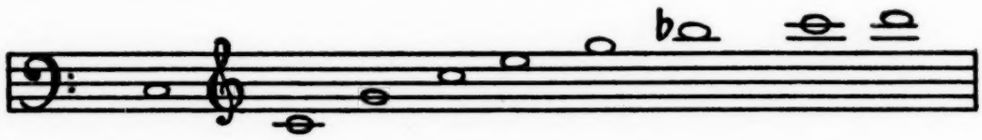
[2] P. Langer, *Die Grundlagen der Psychophysik*, 1876, 79.

[3] G. E. Müller, *Zur Grundlegung der Psychophysik*, 1878, 289.

[4] C. Stumpf, *Tonpsychologie*, 1883, I, 250.

[5] C. Lorenz, Untersuchungen über die Auffassung von Tondistanzen, *Philos. Stud.*, 1891, 6, 26–103.

interpretation of the data.[6] Suffice it to say that the matter was by no means finally disposed of by Lorenz. He established beyond doubt the fact that it is possible for observers, musical and unmusical alike, to make comparisons of supraliminal tone-differences. But the criteria for these comparisons and the possible artifacts peculiar to the method of experimentation which he used were not clarified. The latter point is important. It has been argued [7] that the bisection of a supraliminal tonal distance by the usual method of equal sense distances runs the danger that observers who find it difficult to abstract from musical influence in their judgments are here more prone than ever to allow the aesthetic consciousness to get the upper hand; and that this fact, taken in connection with the well recognized shortcomings of the method of equal sense distances in general,[8] lends support to the doubt regarding Lorenz' conclusion, *viz.*, that the midpoint of a tonal distance falls not at the geometrical, but at the arithmetical mean. If this conclusion were accepted uncritically it would mean that all the adjacent notes on the accompanying staff are separated from one another by equal amounts of sensation since the difference in vibration-frequencies is the same!



Such was the confused state of affairs when casual observation was first supplemented by systematic experimentation. Later work along the same lines by somewhat similar methods has brought with it the appearance of greater order and uniformity.[9] Bisections of intervals smaller than an octave fall uniformly very close to the geometrical mean; but as one goes beyond the octave the midpoint of the interval tends to move up from the geometrical towards the arithmetical mean, progressing from the former to the latter as the size of the interval increases.

It would naturally be supposed that in trying to determine the apparent size of the same interval at different pitches a more direct method of experimentation could be used than the bisection of an interval by the method of equal sense distances. If the question concerns the apparent size of C–G (*e.g.*) as

[6] The reader who is interested in this matter will find an historical and critical discussion of the literature, together with complete references, in Titchener, *Experimental Psychology*, 1905, Vol. II, pt. ii, 232–248.

[7] C. Stumpf, Ueber Vergleichungen von Tondistanzen, *Zsch. f. Psychol.*, 1890, 1, 419 f; H. Münsterberg, *Beiträge zur experimentellen Psychologie*, 1892, IV, 156, 175.

[8] G. E. Müller, *Die Gesichtspunkte und die Tatsachen der psychophysischen Methodik*, 1904, 237 f.; Titchener, *op. cit.*, 203 f.

[9] For references to recent work and an account of new experiments see C. C. Pratt, Bisection of tonal intervals larger than an octave, this JOURNAL, 1928, 11, 17–26.

compared with c–g, the most direct way of finding an answer would be to present to an observer two intervals from different parts of the scale and to ask for a relative judgment as to size. But this procedure has been used in only one or two cases.[10] Stumpf had maintained, on wholly *a priori* grounds, that judgments of tonal distance are especially difficult when no tone is common to the two intervals and still more difficult when the intervals are from different parts of the scale; that the judgment is made easier if the intervals are composed of tones of the same names (c–g and C–G, as contrasted with d–a and C–G); and that if the intervals are too far apart in pitch the observer will transpose one of them down to the compass of the other and then make the comparison.[11] These objections of Stumpf to a frontal attack on interval-comparison, together with the fame which Lorenz' method acquired, have undoubtedly had much to do with the neglect which the more obvious procedure has received at the hands of subsequent investigators.

The only extensive work on tonal distance by a method which required O to make a direct comparison of two intervals of different pitch and without a common tone was reported by Münsterberg in 1892.[12] He sounded four tones, *a*, *b*, *c*, *d*, and instructed O to compare the two distances *a–b* and *c–d*. For most of the experiments, tones *a* and *d* were kept constant, *c* was the variable tone, and *b* was changed from series to series to form, with *a*, a new interval for comparison. The distance between *a* and *d* was about 1000 cents,[13] an interval just a bit larger than the minor seventh. Hence both intervals presented to O for comparison were always within the compass of an octave. The results do not therefore throw much light

[10] Münsterberg, Vergleichung von Tondistanzen, *op. cit.*, 147–177; and O. Abraham & E. M. von Hornbostel, Zur Psychologie der Tondistanz, *Zsch. f. Psychol.*, 1925, 98, 233–249.

[11] Stumpf, *Tonpsychologie*, I, 251 f.

[12] *Op. cit.*

[13] A cent equals an hundredth part of a tempered semitone. For a discussion of the cent and its calculation, see appendix of A. J. Ellis' translation (2d ed., 1885) of Helmholtz' *Die Lehre von den Tonempfindungen, usw.*, p. 437, art. 24, xiii, and pp. 446–457.

on the apparent change in size of an interval with increase in vibration-frequency since the upper interval was never more than a fourth higher than the lower. Even with this small range, however, Münsterberg found a tendency for the higher interval to appear larger as the size of the lower increased. "Das Maximum der Gleichschätzung liegt zwischen dem Punkt des gleichen Schwingungsverhältnisses und dem Punkt der gleichen Schwingungsdifferenz und zwar verschiebt es sich von jenem Punkt zu diesem mit zunehmender Grösse der Normaldistanz." [14]

The experiments to be reported in the present paper were designed to secure more evidence regarding the apparent size of an auditory interval or distance over a fairly wide range of pitch.

The work was done in the Harvard Psychological Laboratory in the spring of 1926. The Os who took part in the experiment were A, Dr. F. J. Pattie; B, Dr. E. G. Wever; and C, Dr. K. E. Zener. All were well trained in psychological observation. C was very musical. A and B, while not unmusical, were unfamiliar with musical intervals and their relationships.

The method of experimentation was a variation and elaboration of the one used by Münsterberg. The Os were presented with four successive tones, *a*, *b*, *c*, and *d*. The interval *a–b* was kept constant throughout the experiments and served as the standard distance in terms of which the comparison-intervals *c–d* were judged.

There were nine comparison-intervals. The first three were taken from a pitch-region about a subfifth (653 cents) higher than the standard; the next three were about a subfifth higher than the first three; and the last three were about the same distance above the second three. The distance between *a* and the highest *d* was 2525 cents, or a little over two octaves. Thus the Os were making comparison-judgments on three groups of intervals which were removed from the standard interval by about a subfifth, a minor ninth, and an acute twelfth respectively. The standard distance, *a–b*, was composed of the tones 204 d.v. and 283 d.v., an interval slightly larger than the fourth (sometimes called the superfluous fourth). The ratio of this interval is 1.3889, or 568 cents. In each of the three groups of comparison-intervals a distance of the same ratio occurred. The other two intervals of each group had ratios of 1.2871 (437 cents) and 1.4978 (699 cents). In other words, there were three comparison-intervals, 437, 568, and 699 cents, repeated at three different pitch-regions to be judged with respect to the standard interval of 568 at the lowest pitch-region. The actual vibration-frequencies of the intervals may be seen in Table I. The three groups of comparison-

[14] *Loc. cit.*, 170.

TABLE I

VIBRATION-FREQUENCIES OF THE INTERVALS USED IN THE PRESENT EXPERIMENT

Standard Interval	Comparison-intervals		
	First group	Second group	Third group
204–283	1. 305–393	4. 440–566	7. 650–837
	2. 295–410	5. 430–597	8. 640–889
	3. 285–427	6. 420–629	9. 630–944

intervals were used as one complete set of stimuli and the nine intervals were arranged and presented to O in haphazard order. The tones which served as stimuli were from a set of four Stern variators [15] with ranges of 200–400, 300–600, 400–800, and 500–1000 d.v. respectively. When the tones were presented in the ascending order the task of O was to judge the distance *c–d* with respect to *a–b*. During half of the experiment the stimuli were presented in the reverse order, in which case O judged the distance *b–a* with respect to *d–c*. Fifty judgments on each of the nine comparison-intervals were secured from each of the three Os.

The instructions to O read:

> "Following the 'ready' signal you will be presented with four successive tones. You are to compare the distance between the third and fourth tones with respect to the distance between the first and second tones. Express your judgment by one of the three categories 'greater,' 'equal,' and 'less.' Occasionally you will be asked to give either a generalized introspective characterization of your judgments on sense distances, or a complete introspection of a given particular judgment."

The Os experienced considerable initial difficulty in following these instructions and felt uncertain about the criteria of their judgments. They were given several practice-

TABLE II

RELATIVE FREQUENCIES OF JUDGMENTS ON TONAL INTERVALS AFTER 'EQUAL' JUDGMENTS HAVE BEEN HALVED BETWEEN 'GREATER' AND 'LESS' JUDGMENTS

The numbers in the first column refer to the comparison-intervals; the size of these intervals in cents is given in parentheses; a = ascending order, d = descending order.

Observers	A		B		C	
	a	d	a	d	a	d
Group I.						
1. (437).....	0.18	0.02	0.04	0.12	0.04	0.10
2. (568).....	0.26	0.14	0.14	0.28	0.32	0.32
3. (699).....	0.62	0.51	0.62	0.66	0.62	0.54
Group II.						
4. (437).....	0.28	0.10	0.22	0.20	0.20	0.12
5. (568).....	0.62	0.22	0.44	0.54	0.48	0.50
6. (699).....	0.84	0.82	0.72	0.82	0.82	0.84
Group III.						
7. (437).....	0.49	0.38	0.42	0.26	0.18	0.46
8. (568).....	0.78	0.78	0.58	0.72	0.66	0.53
9. (699).....	0.92	0.82	0.96	0.82	0.96	0.88

[15] These variators were secured through a grant from the Milton Research Fund of Harvard University.

series before any of their judgments were recorded. After this preliminary training they settled down to the *Aufgabe* without much further difficulty and gave their judgments with a fairly high degree of subjective assurance. When the data were prepared for calculation the 'equal' judgments were halved between the 'greater' and 'less' judgments and the relative frequencies thus obtained subjected to Urban's phi-gamma hypothesis [16] for finding the ideal value of comparison-interval, in each of the three groups, which in 50 per cent. of the cases would appear equal to the standard interval. The frequencies of judgments from which these values were calculated are given in Table II.

The numerical results for the three groups are shown in Table III. These values are given in cents. In reading the table it should be borne in mind that the standard interval was 568 cents at the vibration-frequencies 204–283, which means that, if intervals of the same ratio at different pitches appeared equal to it in size, the average values in the table would be in the region of 568.

Table III

Values, in Cents, of Comparison-Intervals at Three Different Regions of Pitch which Appeared Equal to Standard Interval of 568 Cents

A = ascending order, d = descending order

Observers	A			B			C		
	a	d	average	a	d	average	a	d	average
Group I....	657	697	677	657	641	649	649	673	661
Group II...	529	616	572	590	558	574	568	577	572
Group III..	438	473	455	491	519	505	552	490	521

The average values of the intervals in Group I are all larger than 568 cents. A is 677, B 649, and C 661. The range of the intervals of this group was from 285 to 427 d.v., a pitch-region less than a fifth higher than the standard interval. One would naturally suppose from previous results that two intervals of the same ratio not more than a fifth apart in pitch would appear equal in size. In these experiments, however, the higher of two such intervals appears smaller. In order to be subjectively equal to the lower interval the higher must be increased from 568 cents to about 660 cents. There is no ready explanation for this particular departure from the expected results. It is not impossible that it represents an artifact of the experimental arrangement. At any rate the

[16] F. M. Urban, Hilfstabellen für die Konstanzmethode, *Arch. f. d. ges. Psychol.*, 1912, 24, 236–243.

deviation from equality of ratio is still further in the direction of the geometrical relation of stimulus to apparent size of interval and to this extent is in line with the usual results for intervals of this region.

The intervals of Group II were a little over an octave higher than the standard interval. For all three Os the average values coincide almost exactly with the value of the standard interval, which means that an interval in this region will appear equal to one about an octave lower only when it has the same ratio as the lower interval. This result agrees with previous work done by the method of bisection which has shown uniformly that equal-appearing intervals not exceeding the range of an octave correspond not to equal vibration-differences but to equal vibration-ratios. In the instance before us the ratio of the lower interval was 1.3889 and the ratio of the equal-appearing interval (average of three Os) about a minor ninth higher was 1.3920, an imperceptible difference of 4 cents.

The most significant results are those from Group III. Here the comparison-intervals were a little over a twelfth (octave plus fifth) higher than the standard. Results in the literature have not been conclusive regarding the apparent size of intervals separated from each other by more than an octave. Bisections of supraliminal distances greater than an octave have generally tended to place the midpoint higher than the geometrical mean. In other words, geometrical divisions of such distances cause the upper portions to appear larger than the lower. It is doubtful, however, whether one is justified in drawing an analogy between bisections of large distances and direct comparison of medium-sized intervals separated by large differences in pitch. There is practically no experimental evidence on the latter point. We have only Stumpf's opinion that "das gleiche Interval mit zunehmender Tonhöhe eine zunehmende Distanz in der Empfindung darstellt." It is therefore of especial interest to note the results of the third group of intervals used in the present experiment, for here we have direct comparison of medium-sized intervals separated by over an octave and a half. For

all three Os the comparison-intervals which have the same ratio as the standard appear uniformly larger than the standard: for A, 25 per cent. larger; for B, 12 per cent.; and for C, 9 per cent. The interval which to A is subjectively equal to the standard of 568 cents is 455 cents. Bs interval of subjective equality is 505 cents, and Cs 521. Here we have clear evidence of the increase in apparent size of an interval with increase in vibration-frequency of the components. An interval in the vicinity of c^3 is *not* equal in size to the same interval an octave and a half lower. It appears on the average about half a tone larger.

In any work on tonal distance the question immediately arises as to the degree of coercive power exerted by the musical consciousness in the judgments of O. If two tones, when sounded simultaneously, are taken in a musical setting they possess what has been called *interval-quality;* if perceived under a more strictly psychological attitude (difference in auditory quality between the two components) they possess *distance.* The more usual attitude, except for very non-musical individuals, is the one in which interval-quality is uppermost in consciousness. Under this attitude an interval with the ratio $\frac{2}{3}$ is always a fifth, and possesses the unique characteristic of 'fifthness' no matter from what part of the scale it may be taken. The experience of the present writer, however, has led him to believe that the influence of interval-quality in tonal judgments is not as inescapable as has sometimes been supposed. Except for an occasional O with very considerable musical training, the O in the laboratory who has had some practice in psychological observation finds it not difficult to abstract from the musical aspects of intervals and to concentrate on the difference in auditory quality between the two tones of an interval.[17] An examination of the introspective reports of the Os in the present experiment shows that only one O, C, and he very seldom, was at all bothered by the criterion of interval-quality cutting across that of distance. The reports are all strikingly alike in their reference to the same sort of criterion of judgment. Each tone, as presented, was perceived in more or less isolated

[17] In this connection see Pratt, *op. cit.*, 24 f.

fashion and was placed in its own right, *i.e.*, without reference to its possible musical relationship to the other tones, at a particular point on an auditory qualitative continuum, whence O could then judge as to the relative amounts of separation between the limiting components of the two pairs of points. Each O had his own way of facilitating these judgments of auditory distances. It was not often that the distances were perceived in purely auditory terms. They were supplemented, rather, by various forms of visual and kinaesthetic schematic imagery which served as a sort of framework or point of reference in terms of which the distances could be estimated. This criterion of judgment would seem to correspond very closely to the *Helligkeitsunterschied* which Abraham and von Hornbostel mention in their work on tonal distance.[18]

The intervals used in the present experiment were selected with a view to obliterating musical effect as far as possible. None of them, with one possible exception, coincides with any musical interval. The standard of 568 cents lies about halfway between the fourth and the tritone,—surely not a very 'musical' interval. The smaller comparison-intervals of 437 cents are appreciably larger than major thirds. The larger comparison-intervals of 699 cents are closest to a musical interval, being slightly but not noticeably smaller than tempered fifths. If an O is so unmusical that an ordinary musical interval is rarely perceived with its label of musical interval-quality upon it, surely the distance of a non-musical interval would in all likelihood not be obscured by interval-quality. And the chances are that interval-quality would also tend to disappear to some extent for musical Os when the distances judged correspond with none of the distances from which the characteristics of interval-quality have been built up during musical history. At worst, these non-musical intervals would be perceived as lying somewhere between musical ones rather than as assimilated psychologically to the nearest musical ratios.[19] These facts, together with the Os

[18] *Op. cit.*, 237.

[19] For a discussion of assimilation of non-musical to musical intervals see H. Moran & C. C. Pratt, Variability of judgments on musical intervals, this JOURNAL, 1926, 9, 498 f.

introspective reports, tend to strengthen the writer's conviction that the influence of interval-quality is not ineradicable in work on tonal distance. The quantitative data furnish a still further item of significance in this connection. If interval-quality were the decisive criterion of judgment, then it would follow that the frequencies of judgment on intervals of the same ratio in the three groups would be approximately the same, for musically regarded an interval, say, of 437 cents appears always the same, no matter what the pitches of the components are. That the average frequencies for the three Os for the same intervals in each of the three groups are by no means the same may be seen at a glance in Table IV. Interval 437, *e.g.*, has a frequency of 0.08 in Group I, 0.19 in Group II, and 0.36 in Group III. Surely something has happened to an interval which in one case has a frequency of 0.08 and in another a frequency of over four times that amount. But

TABLE IV

FREQUENCIES OF JUDGMENTS (AVERAGES OF THREE OS) ON EACH OF THE INTERVALS IN EACH OF THE THREE GROUPS

Figures based on Table I

	Group I	Group II	Group III
437 cents	0.08	0.19	0.36
568 cents	0.24	0.47	0.67
699 cents	0.59	0.81	0.89

musically nothing is supposed to happen so far as interval-quality is concerned. It would seem, then, that the Os were not judging on a purely musical basis. For interval 568 the frequencies are nearly tripled from Group I to III. It may be assumed that interval 699, since it is so close to the tempered fifth, has the obvious characteristics of that interval, if it is perceived musically. No musician would admit for a moment that a fifth in the region of c is any different, as interval, from a fifth around c^3. But the Os in this experiment judged otherwise, and the reason must be that they took 699 cents not as an interval but as a distance, for their frequencies of judgment increased from 0.59 to 0.89. Here, then, is additional support for the contention that the Os were not judging on the basis of interval-quality.

The general trend of the results of this experiment may be envisaged more clearly perhaps by noting the diagrams in Chart I. In diagram A the dotted lines represent the change in apparent size of the same interval (568 cents) at different pitches. Diagrams B, C, and D are based on the figures in Table IV and represent the increase in size of each of the comparison intervals as the pitch increases from Group I to III.

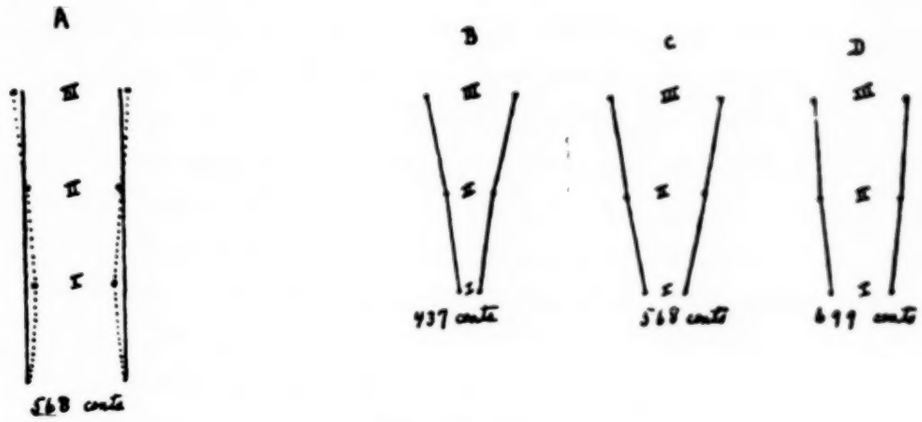


CHART I

SUMMARY

Experiments on the bisection of supraliminal tonal intervals by the method of equal sense distances have shown that the midpoint for intervals not larger than an octave coincides with the geometrical mean, but that in proportion as the intervals exceed the limits of an octave their midpoints tend to move up from the geometrical towards the arithmetical mean. Direct comparisons of the same interval at different pitches, as made in the present experiment, reveal similar phenomena. The interval begins to assume greater size when the pitch is increased by an octave. At an octave and a half its apparent size has increased nearly 100 cents, or about a minor second.

(Manuscript received Sept. 23, 1927)

THE CONCEPT OF THE THRESHOLD AND HEYMANS' LAW OF INHIBITION

BY LLEWELLYN T. SPENCER

Yale University

I. Correlation between the Visual Threshold and Heymans' Coefficient of Inhibition in Binocular Vision

Of the few investigations upon the conditions which determine the stimulus threshold the work of Heymans is peculiarly suggestive. In connection with his experiments on sensory inhibition Heymans (2) found for his single subject that the susceptibility to inhibition which occurred with a stimulus for a given color quality is proportional to the simple threshold for that particular stimulus. His chief conclusion from these experiments is that *the inhibitory power of a stimulus, as measured by the intensity of a stimulus whose effect it can just completely inhibit, is proportional to its intensity.* This law has since been verified for difference thresholds by Anderson (1) and for the simple threshold for albino rats and for human subjects by the present writer (6, 7). The law means that the threshold for a given stimulus is raised in direct proportion to the increase in intensity of a neighboring stimulus which is simultaneously presented.

A quantitative statement of the inhibiting power of the 'inhibitory' stimulus is obtained by calculating the ratio of the liminal rise to the intensity of the inhibiting stimulus. This index is called by Heymans the "coefficient of inhibition." If a given stimulus is subjected to inhibition by stimuli of various compositions, it is clear that the average coefficient of inhibition gives a measure of the susceptibility of the given type of stimulus to inhibition by these other stimuli.

Heymans obtained such measures of susceptibility to inhibition for five different colors and found that they were proportional to the simple thresholds for these colors. For

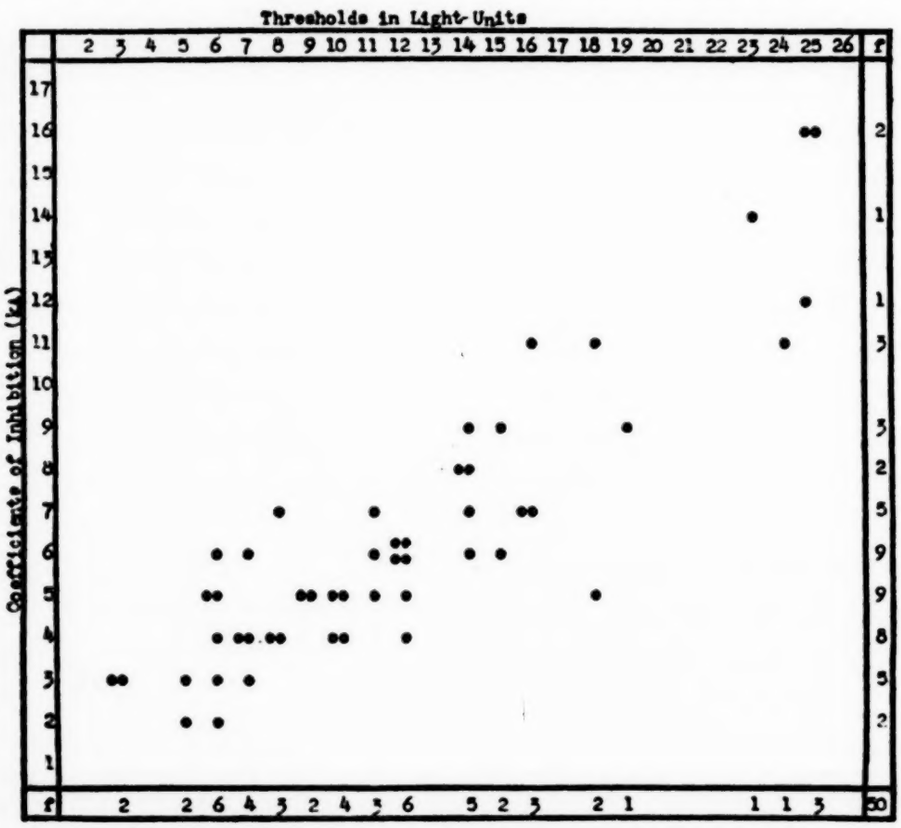
example, the average coefficient of inhibition when red was used as the liminal stimulus bears the same ratio to the simple threshold for red as the average coefficient for green, when used as the liminal stimulus, bears to the simple threshold for green. In other words, the coefficient of inhibition for a given type of stimulus regarded as an inhibited stimulus is correlated with the simple threshold for the stimulus in question. The average figures published by Heymans show a correlation coefficient of + .98, although only five cases are given by him for the correlation. The immediate significance of this correlation is that if a given specific stimulus is particularly susceptible to inhibition the simple threshold for that stimulus is correspondingly high.

The comparison of these measures was made by Heymans for a single subject with rotating discs under daylight conditions, and also in a study of inhibition of gustatory qualities (3). Results obtained from auditory stimuli likewise gave contributory evidence (4). The striking correspondence between these two variables raises the question whether such a correspondence exists, not only among various sensory qualities for a given subject, but also for a given sensory quality among a variety of individuals. An investigation of this question would lead also to information bearing upon another problem connected with Heymans' general law hitherto not fully investigated. This problem is the question whether individuals differ from one another in the susceptibility to inhibition. Heymans reports comparative coefficients for only two subjects. Anderson reports work with eight. The writer's study of inhibition in the albino rat involved only five animals. In all these cases individual variation was found; but a demonstration of variability for a substantial number of cases under similar circumstances is still lacking.

The apparatus and the method used in the present investigation into these two problems (*i.e.* variation in susceptibility to inhibition and the correlation between such variation, if found, and variation in the simple threshold) have been described in a previous paper (7). The increase in the

simple binocular threshold for white light was measured by a photometer (for 50 dark-adapted subjects) as the intensity of a second white light, also binocularly observed, was increased. The increase in the liminal values was found very consistently to be directly proportional to the increase of intensity of the inhibiting stimulus.

The fact of individual variation in susceptibility to inhibition is obvious from an inspection of the data of the present investigation. The correlation-plot given below shows a significant variation in the coefficient which is, as stated above, a direct measure of the individual susceptibility to inhibition. The distribution of these coefficients is heavily skewed toward the upper end, as are also the simple thresholds with which they are to be compared.



SCATTER DIAGRAM FOR CORRELATION BETWEEN t_0 AND k_A; $r = +.894$

The correlation-plot also indicates a decisive answer to the second question; for it shows a high positive correlation (+ .894) between the simple threshold and the coefficient of

inhibition, which is a measure of susceptibility to inhibition. The data may be correlated in three forms. Three thresholds were determined for each subject: (*a*) the simple threshold which we may designate t_0; (*b*) the threshold raised by the simultaneous presentation of a second stimulus A and designated t_A; and (*c*) the threshold raised by the presence of stimulus A' which was twice as intense as A. This threshold is designated $t_{A'}$. We may thus correlate t_0 with the coefficient of inhibition k_A as measured from t_0 to $t_A \left(k_A = \frac{t_A - t_0}{A} \right)$. This treatment gives the measures of the plot shown herewith. We may, secondly, correlate t_0 with the coefficient $k_{A'}$ as measured by the increase from t_0 to $t_{A'} \left(k_{A'} = \frac{t_{A'} - t_0}{A'} \right)$. Finally, we may calculate by the method of least squares the straight line of best fit for the three thresholds—since we have to do with a relationship of the three thresholds presumably approaching rectilinearity—and correlate the best fit t_0 for each individual with his corresponding best-fit coefficient. If every individual had displayed a perfect rectilinear relationship all three of these methods would obviously coincide. Since this is not the case, these three correlations may and do differ. The resulting coefficients of correlation are as follows.

Values Used	r	P.E.$_r$
t_0 and k_A	+ .894	.019
t_0 and $k_{A'}$	+ .919	.015
t_{bf} and k_{bf}	+ .907	.017

The interpretation of these correlations involves the theory of the threshold. Heymans, in view of the correspondence of his figures, concluded that the threshold is due almost exclusively if not wholly to the continuous inhibitory effects of environmental and intraorganic stimuli. He argues as follows (5).

"The fact that as a stimulus is gradually decreased the corresponding sensation is reduced to zero before the stimulus reaches a zero value is usually explained as due to the combined action of two factors: first, that of the resistance which the physiological action of the stimulus has to overcome in the sense organs or in the nervous pathways, and secondly, that of the presence of other, simultaneous stimuli forcing themselves upon the attention. This latter factor falls evidently under the concept of inhibition;

its importance in the occurrence of the phenomenon in question is rendered indubitable by numerous and cogent facts, above all by the decrease in the threshold for light or sound in darkness or silence. On the other hand the hypothetical nature of the former factor will generally be conceded; the assumption is supported, it is true, by certain physical and physiological analogies, but by no direct facts.

As to the relative contribution of each of these two factors to the fact of the threshold, data have hitherto been completely lacking. Moreover, there is little prospect of ever attacking this question through direct experiment. . . . But although a direct experimental appraisal of these two factors is not to be considered, yet one may infer something, it seems to me, in regard to the actual relationship between them from the preceding experimental results. We have found . . . that the thresholds for sensations of different qualities are almost exactly inversely proportional to their resistance to inhibition; that means then that they stand in a constant relation to the increases which they undergo under the influence of an active inhibitory force. This thoroughgoing proportionality between the sensitivity to the effects of inhibition and to the influences which determine the simple threshold points in an unambiguous manner to the conclusion that these latter influences, or at least the preponderant part of them must be of the same nature as the former. . . . If, therefore, it is established from other sources that at least part of the threshold rests upon the effects of inhibition, then we are fully justified in concluding that this part closely approaches the whole. What we are accustomed to measure in our laboratories as the simple threshold must be, ascribed, either wholly or to a degree dependent upon the error of observation, to inhibitory effects which are either incapable of elimination or at least are not eliminated."

The results of the present experiment confirm the relationship which Heymans reports. The following factors may contribute to prevent quite so high a correlation coefficient in the present instance as is obtained from his data. The present values were obtained from single measurements while Heymans' values were obtained after repeated trials. The r given for Heymans' data is calculated from mean thresholds and mean ks. The use of these averages eliminates the extreme variations which might reduce the correlation. Furthermore, the technique of the present investigation did not provide for correction of various experimental errors due to practice, dark adaptation, fatigue, and the like which might considerably reduce the correlation.

A significant fact in regard to the correlation is the failure of the results to exhibit a simple ratio between t_0 and k. While, by a simple ratio, t_0 can be predicted from the coefficient of inhibition, k, with moderate error ($P.E._{est.t} = 1.78$) compared to the observed variation in the threshold ($P.E._t = 3.83$), and the mean threshold, 11.89, a slightly better prediction can be obtained from an equation containing a

constant term. Thus for the data represented by the correlation between t_0 and k_A the regression equation for the prediction of t_0 from k_A is $t_0 = 1.644k_A + 1.67$. The P.E.$_{est.}$ of t_0 from this equation is 1.71 as against 1.78 by the simple ratio. Both of these P.E.$_{est.}$'s were calculated from formulæ.[1] The two other and higher correlations, beside showing a smaller error of estimate, show a lower constant term as well. Heymans' data are also better fitted by an equation containing a constant term, viz, $t_0 = 62.23\ k + .297$, than by the simple ratio of 80 to 1 which he employs. The apparent smallness of this constant term must be considered in relation to the other absolute values of his data, because the constant term of the present investigation, 1.67, is 14 per cent. of the mean value obtained for t_0 while the .297 of Heymans' data is 23.2 per cent. of the mean threshold value of his results.

It may perhaps be possible to explain the presence of these constant terms as due to experimental error, inhering in the apparatus or in the technique of measurement. On the other hand, if the constant terms are true pictures of the facts, *and would persist even when* $r_{tk} = 1$, we are confronted with some constant factor, uniform from subject to subject, which contributes an unvarying influence to the occurrence of the threshold.

The question may now be raised as to whether Heymans' data together with the present correlations justify his theoretical conclusions. Before this question can be answered we must examine the nature of the correlations exhibited and consider what further assumptions are involved. Is the correlation exhibited in both studies spurious or artificial? Does the relation found between t_0 and k proceed from some simple arithmetical or logical implication of k in t_0? The coefficient k is not a simple arithmetic function of t_0 alone, for $k = \frac{t_A - t_0}{A}$ and this ratio contains the independent variable t_A. While this is also a threshold value its amount cannot be

[1] P.E. est. t_0 by ratio $= .67449\sqrt{\sigma_t^2 - 2R\sigma_t\sigma_k r_{tk} + R^2\sigma_k^2}$ when $R = \frac{M_t}{M_k}$.

predicted from t_0 unless k is known. Nor does t_A appear to be a simple logical function of t_0. The simple threshold is usually taken as inversely proportional to sensitivity, and while the same degree of sensitivity that affects t_0 presumably affects t_A correspondingly, it should also affect A. Since A is kept physically constant its effectiveness in raising t_A, if sensitivity is the underlying factor, should vary inversely with variation in t_0, which would produce values for k other than those actually obtained.

The proportionality found between t_0 and k may be stated as indicating that *when t_0 is greater, a smaller amount of an inhibitory stimulus is required to produce a given amount of inhibition, as measured by the increase from t_0 to t_A.* When t_0 is greater, say twice as high, sensitivity is supposedly decreased, in this case cut in half. Hence a physical light-unit has only half the physiological effect when regarded as a visual stimulus. An increase, therefore, of x light-units from t_0 to t_A represents only half as great an increase in "physiological units" as in the case of the original t_0. Half the physiological effect of A is thus required to produce the same objective increase in t_A. With sensitivity cut in half the physiological effect of the physically constant A is actually, by hypothesis, half as great. This would mean that with A physically constant, the increase from t_0 to t_A objectively should remain constant. The results of the investigation indicate, however, that the increases vary directly with the values obtained for t_0. This line of reasoning therefore forces the conclusion that the physiological effectiveness of A, in raising the threshold, does not vary with the physiological effectiveness of the liminal stimulus but produces greater increase in t_0, as expressed by the physical units of the apparatus the higher the values obtained for t_0. This brings us back to the dilemma faced by Heymans. The correlation either indicates that the coefficient of inhibition is merely a function of sensitivity or that sensitivity is a logical function of what the coefficient of inhibition represents, viz, susceptibility to inhibition. It seems difficult to conceive of a variation in sensitivity to the liminal stimulus unaccompanied by a corresponding variation

in sensitivity to the action of the inhibitory stimulus A; whereas, if sensitivity is relatively stable (*e.g.*, maximal) and only susceptibility to inhibition varies, the facts observed are understandable.

Heymans calls attention to the hypothetical nature of the concept of sensitivity contrasted with the indubitable occurrence of inhibition. While the threshold may represent a resistance to excitation, this resistance is at least in part due to the simultaneous presentation of other stimuli and not exclusively to an autonomous inertia of the nervous arc responding to t. In so far as the variation observed in the thresholds of the group or for the various qualities in Heymans' experiment is concerned, this influence of t_0 by k approximates the whole of the influence obtaining. Whatever of autonomous inertia there is may be represented in the constant term referred to; but this is less, on the average, than 25 per cent. of the absolute value of the threshold, diminishes as r approaches 1.00 and, if r true is 1.00 with a term still remaining, is a constant from person to person or from quality to quality.

Whether we accept this reasoning or not, it seems worth while to examine, and if necessary to test by experiment, the further assumptions upon which Heymans' hypothesis is based. His conclusions were based only on the verified occurrence of this relation for qualitative difference within vision, within taste, and within audition. If his theory is true, it must also be true that individual differences in threshold are correlated with susceptibility to inhibition in the sense in question. While this does not follow from his data it is implied in his generalization. The present study brings this to an experimental test with affirmative results. It must also be true, under his hypothesis, that variation in threshold from day to day, if it occurs, must be correlated with variation in susceptibility to inhibition as measured by k. The next paper in this series will be concerned with an experimental investigation of the validity of this assumption. Further, and even more importantly, Heymans assumes that values of k measure susceptibility to inhibition not only in the particular sense investigated, but also in all senses of the given subject

at the given moment. That is, if susceptibility to inhibition is responsible for the occurrence of the visual threshold of a dark-adapted subject, it is inhibition by the stimulation largely of *other* senses which produces the phenomenon. It seems to be Heymans' assumption that the measurement of susceptibility to inhibition by other visual stimuli gives an index of susceptibility to inhibition by stimulation of other and qualitatively different senses. The investigation of the validity of this assumption presents some difficult problems of technique to the experimentalist; but a sufficiently reliable inquiry into its validity seems experimentally possible. If this assumption does prove to be well-founded, it must follow that individual or daily variations in the threshold of one sense are correlated with similar variations in the thresholds of other senses. Consequently the phenomenon of sensitivity or resistance to excitation becomes more than a local phenomenon and assumes the proportions of a physiological characteristic of the organism whose conditional factors are generally and not locally determined.

Bound up with several of these problems are certain related questions. What is the relation of sensory adaptation to both t_0 and k? What effect have variations in efficiency of the sense organs apart from receptivity of the actual nervous tissue? What rôle is played by the general states of well-being or freshness, fatigue, practice, etc.? The phenomena of facilitation, the mechanics of the sense organs, and the limina found for electrical stimulation of muscle-nerve preparations all offer serious obstacles to a ready acceptance of Heymans' view. While these difficulties are not easily disposed of, it seems justifiable to test the various hypotheses above enumerated by experimental methods; for if these fail of verification the larger difficulties cease to require attack, since Heymans' theory would then be no longer tenable.

Summary

Heymans' hypothesis stating that the simple threshold is almost wholly if not exclusively due to the continuous inhibitory influences in the external and intraorganic environ-

ment is described and his data and argument are given. New evidence, concerned with the relation of individual thresholds to individual susceptibility to inhibition, is offered, which shows a correlation coefficient of + .907 between these values. The demonstration of significant variation in susceptibility to inhibition, found in this study, has hitherto been lacking.

The problem presented by the appearance of a constant term in the regression equations is pointed out. No certain evidence is found that these correlations are spurious, but several unverified assumptions are indicated which may be experimentally investigated.

References

1. Anderson, J. E. *Amer. J. Psychol.*, 1925, 36, 461–463.
2. Heymans, G. *Zsch. f. Psychol. u. Physiol. d. Sinnesorg*, 1899, 21, 321–359.
3. Heymans, G. *Ibid.*, 21, 335 f; 346–348.
4. Heymans, G. *Ibid.*, 21, 353–356.
5. Heymans, G. *Ibid.*, 21, 357 f.
6. Spencer, L. T. *J. Comp. Psychol.*, 1923, 3, 389–408.
7. Spencer, L. T. *Amer. J. Psychol.*, 1925, 36, 427–433.

(Manuscript received Aug. 8, 1927)

THE COURSE OF THE AUDITORY THRESHOLD IN THE PRESENCE OF A TONAL BACKGROUND [1]

BY ERNEST GLEN WEVER,

Princeton University

AND

STANLEY R. TRUMAN,

University of California

It is a matter of common observation that certain sounds may come to us upon a variety of auditory backgrounds, and that they then seem to be correspondingly modified. Speech-sounds are perhaps the best case in point, for they occur in all sorts of auditory settings, in silence, in clanging railway coaches, over a telephone that emits a transmission-tone, etc. And it seems very obvious to casual observation that a background of sound interferes with ready hearing; all this is familiar enough. But there has been little systematic study of this phase of auditory perception; just what is the effect of a secondary sound upon hearing seems not to be known.

Perhaps two experiments should be mentioned as bearing upon the problem. Wegel and Lane,[2] following upon the work of Mayer,[3] led two pure tones simultaneously into the ear and confirmed the earlier discovery that the tones interfere with, or 'mask,' one another. If a tone is sounded at such an intensity that it is barely audible, and then a second tone is introduced into the same ear, the first tone can no longer be heard; it is 'masked' by the second, and to be just audible again it must be sounded at a greater intensity; the ratio of this latter intensity to the former is a measure of the masking-effect of the second tone. The masking-effect of a tone increases rather regularly as its intensity is increased. Masking varies also with the relative frequencies of the two tones, though, peculiarly enough, this variation is not always independent of intensity.

A second study in line with our interest is one by Knudsen and Jones [4] in which it is shown that for all persons hearing is reduced in the presence of a noise. These investigators were questioning the ancient belief that so-called paracusics (persons

[1] From the Psychological Laboratory of the University of California.

[2] R. L. Wegel & C. E. Lane, Auditory masking of one pure tone by another, *Phys. Rev.*, 1924, 23, 266–285.

[3] A. M. Mayer, Researches in acoustics, *Phil. Mag.*, (5th ser.), 1876, 2, 500–507; C. Stumpf, *Tonpsychologie*, 1890, 2, 219 ff, 416 ff.

[4] V. O. Knudsen & I. H. Jones, Paracusis, *Laryngoscope*, 1926, 36, 623–663.

with partial deafness due to an impairment of conduction in the middle ear) can hear more acutely in the presence of a secondary sound, and they conducted articulation-tests with a background of noise which showed the belief to be founded upon an illusion. Incidentally they worked with persons of normal hearing also, and found that with a background of noise the threshold of hearing is always increased in comparison with conditions of quiet. This is true whether hearing is tested for speech or for simple tones.

The present investigation somewhat resembles in its setting both the experiments just mentioned, but in important respects it differs from both. In Wegel and Lane's experiment the tones used were simultaneous and (so far as we know) were not sustained; we would therefore be reluctant to regard one as a background for the other. Here, on the contrary, one of the tones is sounded constantly to form a true background of auditory stimulation; that this is an essential point will presently be evident. While our study resembles Knudsen and Jones' experiment in that it involves the testing of (normal) sensitivity in the presence of a constant background of sound, it diverges from that experiment primarily in the use of a tone rather than a noise for the background. And, further, we were not content to base our calculations of the threshold upon one observation, but attempted to take several, and thereby happened upon the discovery that as the background is maintained the threshold does not remain constant, but changes in the direction of greater sensitivity. This investigation thus led itself to the determination of the exact course of the auditory intensive threshold in the presence of a tonal background.

Observers.—The Os were *T* and *W*, the writers of this report, who also served alternately as E. During the experiment they kept separate introspective reports. Their hearing is normal. Two additional Os were called in for one group in Experiment I; these gave results which confirm those from the regular Os.

The work was all carried on at night in the silent-room of the laboratory.

Preliminary Trials

In the preliminary work, which was not very systematic, we employed various devices for generating the 'ground-tone.' Whistles, organ-pipes, and Stern's tone variators, all blown by compressed-air from a tank fitted with a sensitive regulating-valve, afforded us a number of frequencies and

intensities of sound for the background. For the tone for which sensitivity was measured (hereinafter called the 'figural' tone) [5] we used an audio-oscillator of simple type which delivered a tone whose intensity could be regulated by means of a potentiometer.

Procedure.—In the presentation the ground-tone was first turned on, and then the figural tone was given as in the determination of an RL. The Method of Constant Stimuli was first tried; but it was soon abandoned in favor of the Method of Limits by reason of the changing location of the threshold. In this procedure, after the ground-tone was turned on, the threshold for the figural tone was determined in alternately descending and ascending series.

Results.—The threshold in the presence of the ground-tone is at first very high in comparison with the normal threshold under conditions of quiet; but in successive series its value changes until it approaches the normal threshold.

As this change occurs in the threshold there is a striking change in the experience. At first the 'ground-tone' dominates the auditory field, and the 'figural' tone, to be heard at all, must be sounded at a very high intensity. And even then it seems not distinct, but blended with the 'ground-tone.' But a few series later the perception changes, and the 'figural' tone comes out as distinct and separate, with an individuality of its own. As time goes on it grows clearer and richer, standing forth more prominently in the experience, while the 'ground-tone' recedes. In its recession the 'ground-tone' undergoes a striking alteration in quality. At first the tone is dominating and rough, but it grows progressively smoother, simpler, and more mellow. It may even go out altogether, and O then declares in surprise that he has not been 'hearing' the ground-tone at all.

[5] It is convenient to distinguish the two stimulus-tones by use of the terms 'figural' and 'groundal,' irrespective of the psychological experience that is produced in a given case. Typically, as will soon appear, these tones were actually experienced as figural and groundal, respectively, but this was not invariably the case. Whenever, as in our descriptions of the experiences, it seems possible that our use of the terms may lead to confusion, we shall place them within single quotation-marks, or else resort to the more cumbersome expressions, 'the testing-tone' and 'the tuning-fork tone.'

Discussion; the Experience of Figure and Ground.—The experience just described is quite like the more familiar case in the field of vision in which a part differentiates itself from a total visual complex and comes forward as figure, while the remainder recedes to become the ground of the perception. In the examination of one of the familiar picture-puzzles, for example, the first view may give but a confused mass; but soon one feature comes forward and is seen as a face in profile, with its background as the dimly-perceived remainder of the picture. The figure-ground phenomenon in vision has received considerable study,[6] and it has been shown that in general a figure is clear and distinct while its background is vague and ill-defined. The figure has object-character; it is a unit, insistent and self-contained; the ground, on the other hand, is formless, recessive, and lacking in definite boundaries. And according to the results obtained here the auditory case seems essentially the same.

When the 'ground-tone' is first turned on it dominates the auditory field; it is rough in quality, loud and attention-compelling; perceptually it is figure, with a background of silence. And at the first presentation of the testing-tone it, if it appears at all, is blended with the other tone; here the fused tones constitute the figure on a ground of silence. But after a little time the perception changes; what we have called the 'ground-tone' recedes and becomes actually the ground. It loses its preliminary figural character—its roughness, its loudness, its insistence—and falls back to a secondary place in the experience. And at the same time the testing-tone comes forth with strong and definite figural character, with the other tone as its background.

We seem here to have favorable conditions for demonstrating in auditory perception a figure-ground phenomenon essentially similar to that already known in vision. Such a demonstration is of considerable interest from a systematic point of view, for it stands in support of the position assumed

[6] Particularly by E. Rubin, *Visuell wahrgenommene Figuren*, 1921. See also E. G. Wever, Figure and ground in the visual perception of form, *Amer. J. Psychol.*, 1927, 38, 194–226, and titles there cited.

by the *Gestalt* school that the figure-ground relationship is fundamental to perceptual experience.[7]

Experiment I

Since the results obtained in the preliminary work were mainly qualitative in nature, the experiment was continued with more adequate controls to determine for a given set of conditions the exact course of this change in liminal sensitivity in the presence of a ground-tone. Moreover, the normal threshold with silence as a ground was determined for purposes of comparison.

Apparatus.—As a constant source of sound for the background-tone we used a tuning-fork oscillator, which is essentially an electrically-driven tuning-fork of improved design.[8] This instrument delivered to a loud-speaker a tone of constant intensity and frequency (1000 cycles/sec), fairly free from overtones. For the figural tone we used the audio-oscillator mentioned above, with some minor modifications. Its frequency was fixed at 2250 cycles/sec. The intensity could be altered by small steps by means of a potentiometer, and the arrangement of resistances was such that the scale of the potentiometer was practically linear in the region used. The tone was delivered to a watch-case receiver in the silent-room.

O sat alone in the silent-room, 1 meter directly in front of the loud-speaker that emitted the ground-tone. Securely clamped in place over his left ear was the receiver which delivered the figural tone. The receiver did not touch the pinna, but was held in a cup-shaped receptacle that fitted close against the head over the pinna and held the receiver just 6 cm from the orifice of the ear. Perforations in the walls of the receptacle permitted the ready entrance of the ground-tone. The right ear was not plugged, but was covered rather closely by a cap on the right side of the head-piece. A biting-board insured constancy of O's position with respect to the loud-speaker and the walls of the room; when the ground was silence O was not required to use it, but could assume a more comfortable position.

E sat in an adjoining room with all the apparatus except the loud-speaker and the receiver, and there made all necessary adjustments. E and O were in telegraphic connection, and communicated with each other through a simple code.

Procedure.—The procedure in this experiment was rigorously as follows. A click of the relay in the silent-room was O's preparatory signal. E then set the potentiometer to give in the receiver a tone of moderate intensity, and then reduced the intensity continuously until O pressed a key to signify

[7] See esp. K. Koffka, Perception, an introduction to the Gestalt-Theorie, *Psychol. Bull.*, 1922, 19, 553–570; and W. Köhler, *Die physischen Gestalten*, 1920, 183 ff, and Gestaltprobleme und Anfänge einer Gestalttheorie, *Jhrber. ü. d. ges. Physiol.*, 1922, 523 ff.

[8] Halverson has described and figured this instrument. See *Amer. J. Psychol.*, 1927, 38, 294 f.

that the tone had just disappeared. The potentiometer-reading was thereupon recorded. This constituted a D (descending) presentation. Then, 15 seconds from the beginning of this D presentation, an A (ascending) presentation began; here, after the usual signal, E set the potentiometer at zero, and gradually increased the intensity until O signalled the appearance of a tone. The presentation itself took from 7 to 10 sec.; the remainder of the 15 sec. was devoted to recording and to an adjustment of the apparatus. D and A presentations continued alternately every 15 sec. for 5 min., making 20 presentations in all. These 20 presentations, taken in succession with silence as a ground, constitute an *S-series*.

Immediately following this series a new series was run in which the background was the tuning-fork tone from the loud-speaker; this is called a *T-series*. The first presentation (D) began coincidentally with the turning-on of the ground-tone, and A and D presentations continued as before. After the 20th presentation the ground-tone was turned off and O rested or wrote his report. Such an S-series and a T-series constitute a *group*, and three such groups were taken at a single sitting. There were 12 sittings in all, making 36 groups for each O.

Results.—The data for each of these groups indicate the course of the intensive threshold during 5 min. of silence, and then during 5 min. of stimulation by the ground-tone. But as one would expect, such data show considerable variability of the threshold, variability due to uncontrolled changes in the acoustic mechanism, to distractions, etc. Now, since any given order of presentation in a series—say, the 1st, the 2d, or the 20th—is a measure of the threshold after a given period of exposure to the ground, one can compute averages of all the 1st, all the 2d, etc., up to the 20th, presentations, and draw from these final results a curve that can be regarded as typical of the course of the threshold under the conditions, with uncontrolled variations considerably ironed-out in the averaging process. This has been done for each O for both the S-series and the T-series, after discarding the first 6 groups as practice-groups; and these final results are given in the form

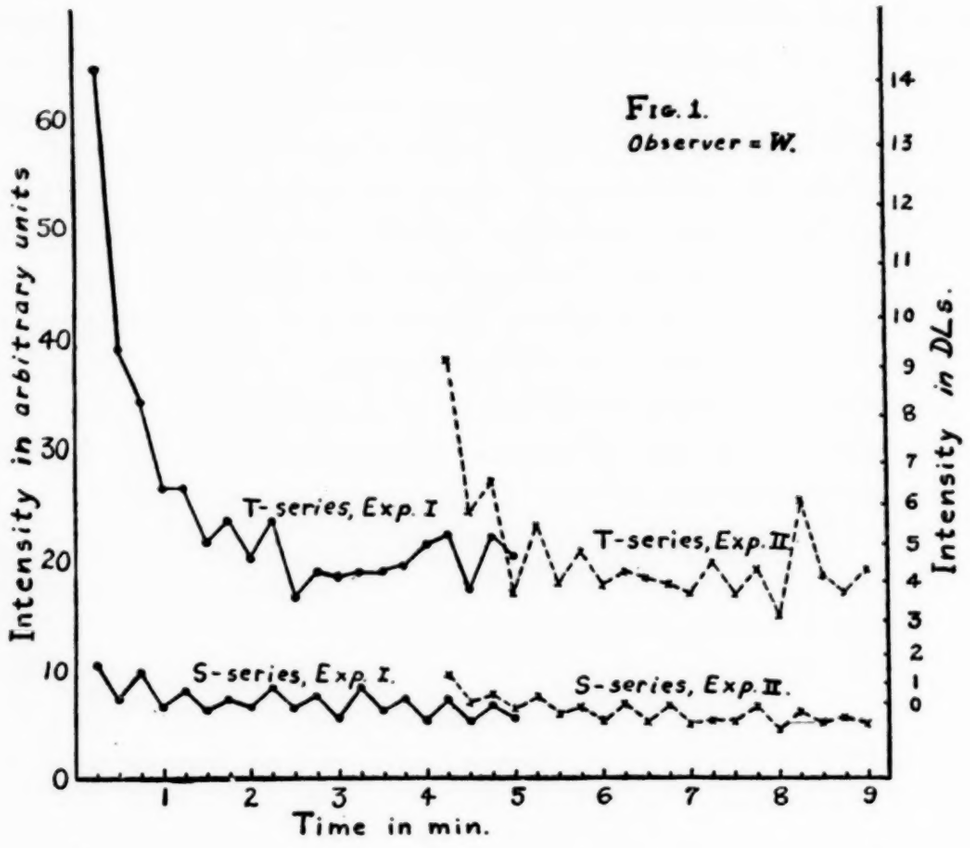


FIG. 1.
Observer = W.

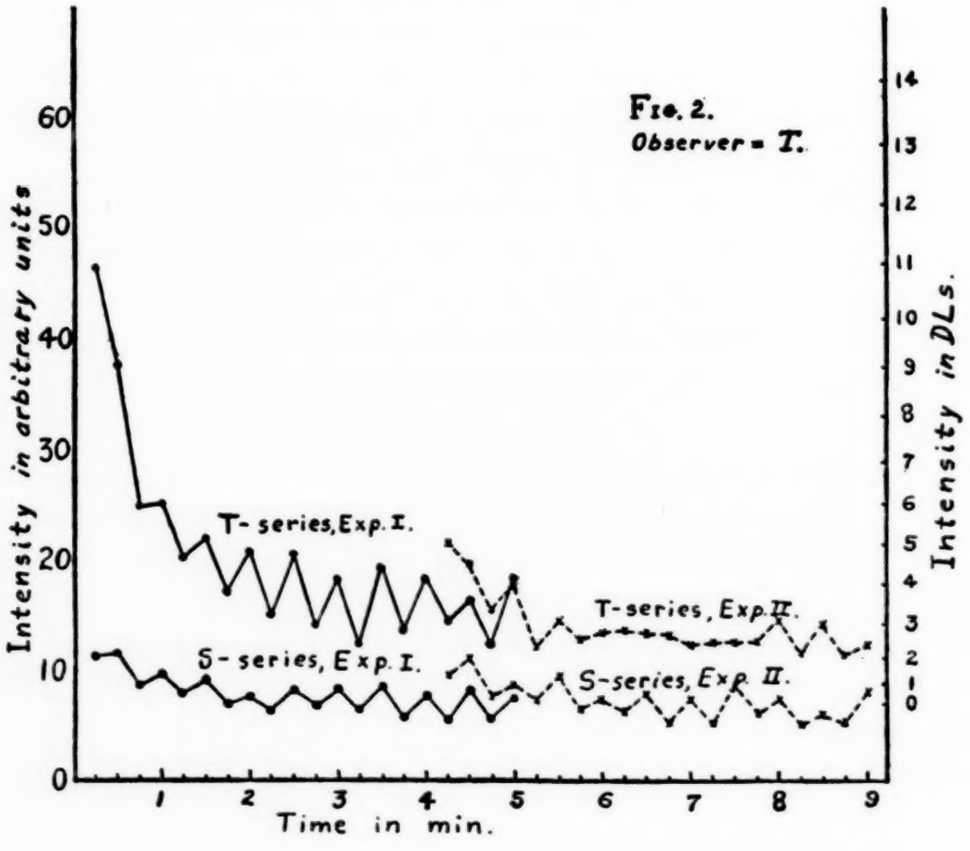


FIG. 2.
Observer = T.

of graphs (the solid-lined curves of Figs. 1 and 2; the dotted curves are to be disregarded for the present). One curve (the lower one in each Fig.) is drawn for the S-series, and another for the T-series. (The peculiarly jagged character of the curves is due to the fact that presentations were alternately D and A, and constant errors of judgment operate so that such D and A observations do not give identical results; this fact does not interfere with the ready tracing of the course of the threshold.) In the Figs. the time is shown on the abscissa, in min., dating from the beginning of the first presentation, while the average threshold-intensity for each presentation is indicated on the ordinate. The left ordinate-scale shows the intensity in terms of the readings of the potentiometer; the units here are practically equal physically, but of course arbitrary. But by determining empirically the DLs for one of the Os (*T*), at various regions of the scale of the potentiometer, we have calibrated the instrument in sensation-units; the right ordinate-scale thus gives the ordinate values in DLs.

From the Figs. it is apparent that during the 5 min. of testing in the S-series the threshold has remained fairly constant; the slight downward slant of the curve is not here significant.[9] But the curves for the T-series show a remarkable change in the threshold during the period. For W (Fig. 1) the curve when the ground-tone is first turned on is over 14 DLs above the normal; after about 2 min., however, it has swept downward until it is but a little over 4 DLs above normal. It has then practically reached its minimum, and drops no more than does the normal curve. For *T* the case is the same except for the extent of change; for this O the threshold begins at about 11 DLs and in 2 min. has become about 4 DLs, thereafter to remain fairly constant.

Discussion; the Question of 'Masking.'—The above results have shown a progressive shift in the intensive threshold after

[9] All the curves, with perhaps one exception, show this downward slant, which we have concluded from examination of the results from sitting to sitting to be the effect of general practice in detecting the presence of a faint sound. Though the Os were highly practiced to begin with, they displayed nevertheless a gradual improvement throughout the experiment; the curves given depict a downward tendency because they are in effect a composite of successive segments of this longer practice-curve.

the introduction of a tonal background into the auditory field. At first the threshold is considerably raised, but recovery of sensitivity is rapid, until (under our conditions) after two minutes or so it is only about 4 DLs above its normal value. This shift of the threshold is of crucial significance in the question of the masking of one tone by another. In the experiment of Wegel and Lane, referred to above,[10] the setting resembled that at the beginning of our T-series. These experimenters led into the ear a pair of tones, and raised one of them to bare audibility in order to determine the threshold. They made four determinations in succession for each pair, and averaged them in plotting their curves. How well these determinations agreed they do not relate. Nor do they state what precautions (if any) were taken to insure constancy of the temporal relations, though, as our results show plainly, the temporal relations in such determinations are of first importance.

Perhaps their measure of the masking-effect is a measure of the threshold before it has proceeded very far in the course of change shown in our curves; in this case we must sound the caution that the masking-effects as shown by them probably do not apply when any one of the tones is sustained. If, on the other hand, their procedure involved so much delay after the sounding of one of the tones that their determination of the threshold fell on the tail of the curve of sensitivity, then we should insist that their data are probably not applicable to the alternative case of occasional simultaneous sounds. Certainly we must doubt that the values for 'masking' which they have so carefully worked out can be expected to hold for any but a single temporal condition.

Wegel and Lane's results have been applied especially to the problem of the interference of concomitant tones in telephony. But since, as is well known, interfering sounds in the telephone are of various sorts from momentary to continuous, it is evident in the light of the above discussion that such an application must be made with great caution. A real solution of the problem of telephonic interference must

[10] Cf. footnote 2.

await the determination of 'masking' as a function of time as well as of pitch and intensity.

As has already been mentioned in the account of our preliminary series,[11] the change in sensitivity shown in the curves above is paralled by a change in the observer's experience. And we have regarded this change as a shifting of figure-ground relationships within the perception; more explicitly, as an emergence of the one tone as figure and the recession of the other to become the ground. Why such an alteration of figure-ground relations should occur in this situation is an intriguing question. We sought to throw some light on the matter by first discovering whether the basis for the change was to be found in the psychological process of rendering a judgment upon the auditory complex, or whether it was something more 'automatic' (as, *e.g.*, sensory fatigue) and quite independent of the judgmental situation. Accordingly, we devised a second experiment in which the ground was introduced some time before the judgments began, and during which preliminary time the attention of the O was kept occupied with some suitable task. It is obvious, now, that if the change which occurs in the threshold is a function of the rendering of judgments (attentional analysis of the auditory complex, *e.g.*), the curve that is obtained in this second case after judgments are begun should be the counterpart of the one obtained in the experiment just described; *i.e.*, since no judgments have been made and attention has (presumably) been otherwise occupied, the preliminary maintenance of the ground should have no effect. But, in another case, if the effect of the ground is mechanical and proceeds quite apart from the making of judgments, as, *e.g.*, sensory fatigue would be expected to do, then the situation should have changed during the preliminary interval of stimulation, and the sensitivity of the ear should, after a given time, be found to be at about the level that would have been shown had one been testing it continuously as in the first experiment. And if, under our conditions, the preliminary period were over 2 min., the curve obtained after testing was

[11] P. 100, *supra*.

begun should be practically a horizontal line from the beginning, running 3 or 4 DLs above the normal curve for a ground of silence. A determination of the curves under these conditions is the object of our next experiment.

Experiment II

Apparatus and Procedure. Actually, this experiment was performed simultaneously with Experiment I, with the sittings of the two intermingled to equalize any practice-effects and thus to make the results directly comparable.

The same apparatus was used and the procedure differed only in that the ground—whether silence in the S-series or the tone in the T-series—was maintained for 4 min before the presentations began. That is, in the S-series O sat in the silent-room, engrossed in reading or in mathematical calculation, for 4 min before the first presentation; thereafter the series went on just as in Experiment I for 20 presentations. In the T-series the ground-tone was turned on and continued; during the first 4 min. O read or calculated, and then the presentations began; after the 20th presentation the ground-tone was turned off. Such an S-series and T-series together constitute a *group*, and after each group O and E changed places. Three groups were taken for each O at each sitting. There were in all 12 groups for each O, of which the first 2 were called practice-groups, and the data from the last 10 averaged for drawing the curves.

Results.—The averaged results are shown in the dotted curves of Figs. 1 and 2. The curves begin at the 4-min. line and continue for 5 min.

It is seen that in all the series (S-series and T-series alike) the curves begin above the level that would have been attained had judgments been made during the 4-min. interval. In the S-series the difference is not very great. In the T-series the curves begin somewhat below the level reached after 4 min. in Experiment I, and drop rapidly until after $\frac{3}{4}$ min. more they have about reached that level which they then maintain throughout the remaining presentations.

On the experiential side the T-series, under our conditions, showed important differences from the first experiment. No longer did the two tones blend in the early presentations, but from the beginning of the testing-period the receiver-tone stood out with figural character. The tuning-fork tone had already taken on the mellow, attenuated character that it was found to assume later in the series in Experiment I. Apparently, then, the preliminary sounding of the tuning-fork

tone sufficed to establish it as ground, and when the new tone appeared in the perceptual situation it at once assumed figural status.

Discussion; the Cause of the Change in Sensitivity.—The subjective difference between this and the first experiment correlates well with the results shown in the curves. From the beginning of the testing-period the threshold is considerably lower than it is at first in Experiment I; this initial difference is about 5 DLs for W and 6 DLs for T. From this we infer that the preliminary sounding of the 'ground-tone' has in some automatic way (*i.e.* aside from the rendering of judgments) resulted in a considerable increase of sensitivity to the 'figural' tone.

Yet this increase has been limited in extent; the curves are not essentially straight lines, but begin above the minimal level of Experiment I and quickly sweep downward to that level. This sweep of the curves is manifestly a function of the rendering of judgments; it would seem that during the first 3 or 4 presentations there is a kind of "warming-up" in the process of judging the presence of the testing-tone in this situation.

Just what is the 'automatic' process resulting from the preliminary sounding of the ground-tone we are not prepared to state. The case may be merely that the new feature of the situation (the testing-tone) by its very novelty obtains an attentional, and hence an intensive, advantage over the maintained and now monotonous 'ground,' and that such advantage is cumulative, up to a certain point, from presentation to presentation. There is good authority for the supposition that attending to a stimulus renders it more intense.[12] Certainly, the testing-tone caught the attention more readily in this second experiment; this was shown in the reports of the Os. Moreover, it will be recalled that here the testing-tone assumed figural status at once; that figurehood

[12] As, *e.g.*, O. Külpe, who in his *Outlines of Psychology* (Engl. trans.), 1895, p. 429, maintains that "the sensation of a loud sound, inattentively experienced, may seem equal . . . to that of a faint sound, attentively experienced." For a full discussion of the relation between attention and intensity, see E. B. Titchener, *The psychology of feeling and attention*, 1908, 212 ff.

and focality in attention run pretty much hand-in-hand has already been demonstrated in visual perception.[13] And, further, in visual experience it has been shown [14] that figural potency holds over from one presentation to the next; this is called the 'figural after-effect,' and we can very well suppose that such an effect is cumulative in character, and can operate in auditory perception as well as in visual to enhance the figural status of a stimulus in successive presentations. Now, if figural, or attentive, status be translated into increasing intensity, we have a tenable hypothesis to account for the initial increase in sensitivity in Experiment II, and for a large share of the change in sensitivity shown in the curves of Experiment I.

An alternative, and rather fascinating, hypothesis is that the process is one of sensory fatigue.[15] It is conceivable that the ear becomes progressively fatigued to the ground-tone, and that thereby the masking-effect of that tone is gradually removed, leaving the organ increasingly sensitive to the testing-tone. Now, to make this explanation of our phenomenon plausible, we should have to show auditory fatigue to be progressive, to attain a maximum (under our conditions) in about 2 min., and to be specific, *i.e.* selective to the fatiguing tone. But, unfortunately, the present state of our knowledge —or perhaps we should say, our ignorance—of auditory fatigue is one that lends little encouragement here even to speculation. Weinberg and Allen [16] think that fatigue becomes maximal in about 2 min.; Flügel [17] got maximal fatigue (usually) in 40 sec.; while Pattie's [18] results are uncertain; it seems probable that the intensity of the stimulus makes a

[13] E. G. Wever, Attention and clearness in the perception of figure and ground, *Amer. J. Psychol.*, 1928, 40, 51–74.

[14] Rubin, *Visuell wahrgenommene Figuren*, 1921, 6–21.

[15] These two hypotheses are not mutually exclusive; perhaps in this situation the novel stimulus has attentive value because, through fatigue, sensitivity has become dulled to the other tone.

[16] M. Weinberg & F. Allen, On the effect of aural fatigue upon the critical frequency of pulsation of tones, *Phil. Mag.*, (6th ser.), 1924, 47, 127.

[17] J. C. Flügel, On local fatigue in the auditory system, 1920, 11, 128 ff.

[18] F. A. Pattie, Fatigue in the auditory mechanism, *Amer. J. Psychol.*, 1927, 38, 52 f.

difference in the development of fatigue. Weinberg and Allen got high specificity of the fatigue-effect under their conditions; Pattie found only general, non-specific fatigue. It is true that the two experiments were widely different in their settings, but if both were dealing with the same thing one would expect a certain accord. Pattie's results might even be construed as casting doubt on the *progressive* character of auditory fatigue.[19] This field is all uncertainty, and there seems little to be reaped from it for our present uses; we venturously suggest that perhaps the situation may be reversed, and if somehow it can be established that our curves are the result of auditory fatigue, then in the procedure of this experiment we have a simple and reliable method that may be applied to the measurement of the extent, course, and duration of fatigue in the auditory mechanism. It is this possibility that gives the latter hypothesis its fascination.

It would appear—though we take the results of this analysis as suggestive only—that the change in liminal sensitivity revealed after the introduction of a tonal background is due in a measure to judgmental factors, but primarily to something or some process more 'automatic' in character, such as novelty or sensory fatigue.

Summary

1. In this experiment it has been shown that the auditory intensive threshold, as measured in terms of a particular tone, is considerably raised at the introduction into the auditory field of a second, constant tone.

2. Under our conditions, however, the threshold does not remain at its high point, but immediately begins to decline. The recovery of sensitivity proceeds at an ever-diminishing rate until after about 2 min. it has practically ceased. From then on the curve is essentially level, even up to 9 min. of stimulation by the ground-tone, with no indication of further change. Within this time-limit (9 min.) the average threshold is never so low as when the background is silence.

[19] Weinberg & Allen, *op. cit.*, 130 ff; Pattie, *op. cit.*, 51 f, 54; see also Flügel, *op. cit.*, 117 f, 131 f.

3. The lowering of the threshold has been shown to be in part a function of the judgmental situation, but due primarily to some factor that appears to operate quite automatically. It is suggested that this factor is an enhancement of the intensity of the 'figural' tone through attention, or else the progressive removal of the masking-effect of the 'ground-tone' through auditory fatigue.

4. Correlated with the change of threshold has been found a change in the observer's experience, a change which is best characterized as the emergence of one tone as figure and the recession of the other to become the ground of the perception. The evidence is in favor of the view that the figure-ground relationship holds in audition as in vision, and is probably a general characteristic of perceptual experience.

(Manuscript received Oct. 3, 1927)

THE RELATION OF PHYSICAL CONSTITUTION TO GENERAL INTELLIGENCE, SOCIAL INTELLIGENCE AND EMOTIONAL INSTABILITY

BY H. E. GARRETT AND W. N. KELLOGG

Columbia University

The subjects used in this study were 221 Columbia College freshmen, all presumably normal individuals. As a measure of physical constitution a simplified form of the Naccarati "morphologic index" was used, adapted so that it could be computed from measurements made on photographs of our subjects taken in the university gymnasium. The Thorndike Intelligence Examination for High School Graduates, the George Washington Social Intelligence Test, and the Woodworth Personal Data Sheet were employed as measures of general intelligence, social intelligence, and emotional instability, respectively.[1]

A brief review of studies relevant to the present report will be considered before the experiment itself is described. No attempt will be made to cover all of the research on the relation of physical and mental traits, as those interested will find fairly complete historical summaries in the monographs of Sommerville [2] and Wertheimer and Hesketh [3] and elsewhere.

Physique and General Intelligence. In 1905, Viola, following the anthropometric studies of De Giovanni, differentiated three morphologic types which he called microsplanchnic, normosplanchnic, and macrosplanchnic.[4] Microsplanchnics, according to this author, possess relatively small trunks and long arms and legs; macrosplanchnics, large trunks and short arms and legs; and normosplanchnics, who fall in between the two extremes, show an evenly balanced development of both trunk and limbs. As a convenient single measure of body-type, and for purposes of correlation with other measures, Naccarati, in his study of physique and intelligence, devised a "morphologic

[1] This study is one of several supported by a grant from the Council for Research in the Social Sciences, Columbia University.

[2] R. C. Sommerville, Physical, motor, and sensory traits, *Arch. of Psychol.*, 1924, 75.

[3] Wertheimer & Hesketh, The significance of physical constitution in mental disease, *Med. Monog.*, 1926, vol. 10.

[4] S. Naccarati, The morphologic aspect of intelligence, *Arch. of Psychol.*, 1921, 45.

index" (M.I.), essentially the ratio of the length of the limbs to the volume of the trunk. High M.I.s are, of course, associated with microsplanchny, low M.I.s with macrosplanchny, and intermediate values with normosplanchny.

The calculation of the morphologic index offers several drawbacks where large groups are concerned, because of the large number of measurements needed, and the expert knowledge of anatomical points required.[5] Hence, as a simpler measure of body-type, Naccarati has generally used the ratio Ht/Wt, which index possesses the advantage of being very much easier to secure than the M.I., and which can, moreover, be readily employed with large groups. The correlation between M.I. and Ht/Wt was found to be uniformily high; in one group of 50 subjects, r was .75, while in another group of 75, r was .70.

Measurements for determining body-type should be taken in youth, preferably between the ages of 20 and 25 years. This is due to the fact that bodily growth in childhood is too rapid for reliable determination; while after 25 years occupation, disease, the tendency to gain in weight, etc., coöperate to change the M.I., usually in the direction of macrosplanchny. Because of these considerations, in his study of morphology and intelligence, Naccarati used college students with an age range from about 17 to 22 years. In this group, the correlation between Ht/Wt and general intelligence (Thorndike score) was .228 ($N = 221$) and the correlation between M.I. and Thorndike score was .356 ($N = 75$). On the basis of these results, Naccarati held that there is evidence of a small—but reliable—positive relationship between general intelligence and body-type.

Later work done by Naccarati in collaboration with others, while it does not positively confirm, at least does not entirely discredit, these earlier conclusions. In one study, r's of $.44 \pm .06$ between Army Alpha and Ht/Wt ($N = 50$) and $.14 \pm .07$ between Otis Advanced Examination and Ht/Wt ($N = 94$) were found, while in a large group of 252 college students, the r between general intelligence (Thorndike) and Ht/Wt was $.13 \pm .04$.[6] In another study there was some evidence that college men of high M.I. tend to surpass those of lower M.I. in such functions as digit span, coördination, and speed of controlled association.[7] In this study, however, the groups were small and selected ($N = 57$) and the range of average differences, though fairly constant, is too small to be statistically reliable. In the other study mentioned, the small size of the group, in one case, and the small and unreliable r, in the other, leave the existence of significant relationship between intelligence and Ht/Wt open to question.

Results favorable to Naccarati's thesis have been reported by Wood[8] and by Mohr and Gundlach.[9] Using Ht/Wt, the Naccarati M.I. and the Thorndike Intelligence Examination (Part I) as his variables, Wood obtained a correlation of .252 between Thorndike and M.I. ($N = 67$) and .235 between Thorndike and Ht/Wt ($N = 192$). Correlations between Ht/Wt and M.I. and parts of the test ranged from $-.023$ to .275. The r between the first semester grades and M.I. was .123 ($N = 67$). These correlations are statistically significant; but they are too low for prediction. In general, they confirm the results obtained by Naccarati. They indicate, too, that Ht/Wt is apparently as good an index of mental ability as the more elaborate M.I.

[5] Naccarati, *op. cit.*, p. 5.

[6] S. Naccarati & R. Guinzburg, Hormones and intelligence, *J. Appl. Psychol.*, 1926, 6, 221–234.

[7] S. Naccarati & H. E. Garrett, The influence of constitutional factors on behavior, this *Journal*, 1923, 6, 455–465.

[8] B. D. Wood, *Measurement in higher education*, 1923, 46.

[9] G. J. Mohr & R. H. Gundlach, this *Journal*, 1927, 10, 117–157.

Mohr and Gundlach's experiments were carried out on 89 male convicts (selected from about 600) of whom 44 were pyknic (roughly macrosplanchnic) in type, and 45 were asthenic and athletic (roughly microsplanchnic and normosplanchnic). Using an index of build obtained by dividing the sum of the circumferences of chest, abdomen, and hip plus weight, by height, they report an r of $-.34$ between this measure and Army Alpha.[10] These investigators also found reliable differences in speed of writing, reaction time, cancellation, and information in favor of the asthenic-athletic group as against the pyknic. While expressly disclaimed by the authors, it seems highly probable that age differences play an important rôle in these results. The pyknics averaged 6 years older than the asthenics and athletics, 11 are above 45 and 1 above 55, while 17 asthenics-athletics are below 25 as against 10 pyknics. Army Alpha is known to show decreases with age,[11] and it is probable that some of the other functions do also. The authors mention lack of coöperation and often open hostility in some of their subjects, and these facts should also be considered in interpreting their results.

Two recent studies, one by Heidbreder [12] and the other by Sheldon,[13] report results opposed to those cited above. Heidbreder computed the correlations between the Ht/Wt ratios of 500 men and 500 women students at the University of Minnesota (taking the two groups separately) and their scores on five different mental tests as well as their final "general intelligence" scores on the composite of the five tests. The r between final score (intelligence) and Ht/Wt was .03 for the men and .04 for the women. The highest r (.10) was obtained between Vocabulary and Ht/Wt for the 500 men. All of the correlations are insignificant in size, and this coupled with their unreliability and the relatively large samples seems definitely to answer in the negative the question of the relation of Ht/Wt ratio to intelligence measures of the kind generally used. Heidbreder, to be sure, did not use the Naccarati M.I. in her work; but the existence of a high correlation between M.I. and Ht/Wt serves to cast considerable suspicion upon the value of this more elaborate measure of body-type.

The work of Sheldon seems to indicate that this suspicion is well founded. This investigator computed the M.I.'s (using Naccarati's technique) of 450 Chicago University freshmen, 17 to 22 years of age. The r between M.I. and Intelligence [14] was $.136 \pm .03$, the r's between the separate parts of the intelligence test (nine in all) and M.I. ranging from $-.023$ to .095. Sheldon interprets this last result to mean that no single measure can be responsible for the relation between body-index and intelligence scores. There was no evidence, according to the author, of curvilinearity of regression in any of the tables.

A summary of the results of these investigators seems to indicate clearly that there is little, if any, relationship between Ht/Wt and measures of intelligence, and that the relationship between general intelligence and the Naccarati M.I., while consistently low, is probably reliable.

Physique and Temperament. The establishment of a relation between the body-

[10] The negative sign is due to the fact that the index here used is the reciprocal of that used by Naccarati. The writers erroneously state (page 134, footnote) that Naccarati found a correlation of .228 between weight/height and Thorndike score. The index should be Ht/Wt.

[11] *Mem., Nat. Acad. Sci.*, 1921, 814 ff.

[12] Intelligence and height-weight ratio, *J. Appl. Psychol.*, 1926, 10, 52.

[13] Morphologic types and mental ability, *J. Pers. Res.*, 1927, 5, 11.

[14] The psychological examination published by the American Council on Education, 1924, was employed as the measure of general intelligence.

type and the temperamental or emotional make-up of the individual seems to offer more in the way of definite promise than the establishment of a relation between body-type and general intelligence. Most of the work in this field has been done by psychiatrists primarily interested in the significance of the physical constitution in mental abnormalities. The best known classification of body-type from this point of view is that of Kretchmer[15] who has distinguished four somatic types; the asthenic, the athletic, the pyknic, and the mixed or dysplastic. These are purely morphological types, the asthenic with his long extremities and small, cylindrical trunk corresponding closely to the microsplanchnic of Viola; the more symmetrically developed athletic resembling the normosplanchnic; and the pyknic with his short legs, thick neck, and large barrel-shaped truck resembling Viola's macrosplanchnic. The dysplastics include those not fitting well into any one of the other groups; the eunochoid is an example of this mixed type.

Of chief importance to Kretchmer is the identification of his morphological types with various syndromes of mental disease. From his clinical studies, he concludes that circular or manic-depressive insanity is generally associated with the pyknic type, and that schizophrenia or dementia præcox is usually found in the asthenic, less often in the athletic and dysplastic types. Kretchmer's observations upon which these conclusions are based were carried out upon 260 Swabian folk of whom 85 were manic-depressives and 175 præcox dements. Among the 175 DPs only two were pyknic, 70 per cent. of the remainder being asthenic and athletic, and the rest unclassifiable. Of the 85 MDs, all but 13 (4 asthenics, 3 athletics, and 6 mixed) were pyknic or pyknoid.

Among normal people in general, Kretchmer distinguishes two large groups or 'personality-types,' the schizothymes and the cyclothymes. The schizothymes grade over into borderline individuals of schizoid temperament, and these, in turn, into the schizophrenics. In the cyclothymic group, the progression is from cyclothymic normal to cycloid temperament to circular insanity. Individuals of cycloid temperament, according to Kretchmer, tend to be sociable, good-natured, cheerful, easily excited or depressed; those of schizoid temperament are reserved, unsociable, serious, and oversensitive. To a lesser degree these descriptions characterize the "normal" schizothymes and cyclothymes. The latter are generally of pyknic or pyknoid type, the former of asthenic or asthenic-athletic type.

A recent study of Kretchmer's constitutional types has been made by Wertheimer and Hesketh.[16] Sixty-five patients (males) from the wards of the Henry Phipps Psychiatric Clinic of the Johns Hopkins Hospital were observed and measured. The group contained 19 manic-depressives, 25 schizophrenics, 11 psychopathic personalities and psychoneurotics, and 10 organic-reaction types. After considering and discarding several morphologic indices previously used in anthropometry, the authors developed[17] the following index of their own:

$$\text{Mor. Index} = \frac{\text{leg length} \times 10^3}{\text{transverse chest diam.} \times \text{sagittal chest diam.} \times \text{trunk height}} \times 100.$$

Using this formula the indices of the 65 patients were computed. The results for the clear manic-depressives and the clear schizophrenics were as follows:

	Av. Index	Range
All Clear MD (11)	233.2	154–318
All Clear DP (23)	280.7	178–337

[15] *Physique and character*, New York, 1925.
[16] *Op. cit.*, 16–72.

Since the age factor (as previously indicated) has considerable influence on morphologic type, the following results for all cases above (and including) 29 years are given:

	Av. Index	Range
Clear MD (6)	199.1	154–262
Clear DP (12)	260.3	178–315

Of the clear MDs at all ages, 45.5 per cent. were pyknic, 36.4 per cent pyknoid; of the clear DPs 78 per cent. were asthenic, asthenic-athletic, or mixed. Of the clear MDs 29 years old and over, 100 per cent. were pyknic or pyknoid, while of the clear DPs, 29 years old and over, 58.4 per cent. were asthenic, athletic, or asthenic-athletic mixed.

The average index values and the range of each of the three main morphologic types are given as follows:

	Index	Range
Asthenic	309	282–337
Athletic	278	248–306
Pyknic	192	153–244

It will be noted that there is no overlapping in the two extreme groups, but a gradual progression from low to high indices. From their statistical distributions as well as from clinical observation, the authors reach the conclusion that there are no clear-cut morphologic types: that the most frequently recurring type is the mixed; the pyknic and the asthenic giving the impression of types because they lie at the extremes of the morphologic distribution. While these two extreme constitutions can be differentiated by their morphologic indices, the differentiation of either the athletic or the mixed from either extreme is considerably less certain.

This conclusion as to the absence of clear-cut morphologic types was also reached by Mohr and Gundlach.[18] Their distributions both for estimated types and physical measurements indicate a regular progression from asthenic at one end of the scale to pyknic at the other. Differences in speed of writing, speed of reaction time, cancellation, etc., in favor of the asthenic-athletic group, and in learning and attention in favor of the pyknics are interpreted by the authors as meaning that the asthenics are more schizoid, and the pyknics more cycloid in temperament. Age differences and the differences in racial ancestry between the two groups may account, again, for part of the differences found.

Naccarati has also attempted to classify psychoneurotics on the basis of morphologic index.[19] His group consisted of 100 Italian male psychoneurotics, ranging in age from 25 to 40 years. One half of the group had been diagnosed as suffering from "asthenic psychoneurosis"—practically equivalent in Naccarati's terminology to neurasthenia, and the other half from "emotional psychoneuroses," *i.e.* hysteria, anxiety neurosis, traumatic neurosis, etc. Naccarati gives the average M.I. of his "asthenics" as 456.64, and of his "emotional psychoneurotics" as 362.06. From this difference of 94.58, he concludes that neurasthenia is more often found among microsplanchnics; hysteria, anxiety, etc., among macrosplanchnics. Naccarati gives no

[17] A minimum scheme for the study of the morphologic constitution in psychiatry, *Arch. Neur. & Psychiat.*, 1927, 17, 93–98.

[18] *Op. cit.*, 154.

[19] The morphologic basis of the psychoneuroses, *Amer. J. Psychiat.*, 1924, 3, 527–544.

measure of variability for either of his averages. SDs calculated by the writers from his data were found to be 88.01 for the "asthenic" average, and 100.21 for the "emotional psychoneurotic" average. This indicates considerable overlapping in the two groups, though the difference between the means is statistically reliable.

An attempt to classify the functional psychoses on the basis of the Naccarati M.I. has been made by F. C. Shaw.[20] One hundred patients were the subjects of the study, 50 were manic-depressives and 50 were dementia præcox, each group equally divided into men and women. The average M.I. of the DPs was 572.3; of the MDs, 402.4, a difference of 169.9 points in index. Shaw considers this difference somewhat affected by the fact that the DPs were on the average 8 years younger than the MDs. The difference in average M.I. for the women alone (the average age difference here was only one year) was, however, 212.6 points in favor of the DPs. All of the patients with M.I.s between 250 and 340 were MDs, all with M.I.s between 680 and 880 DPs. The bimodality of his distribution of M.I.s (the modes were at 400 and 570) offers further evidence, according to this author, for the existence of two types. No measures of variability are given for the averages, so that one is unable to tell just how reliable the obtained differences are.

One further study may be cited. Naccarati and Garrett[21] have studied the relation between M.I. and "emotional stability" as measured by the Woodworth Personal Data sheet, using as subjects 54 college men. Their results indicated that men with low M.I.s tend to answer more questions "wrongly," *i.e.* to be more emotionally instable, than those with relatively high M.I.s. But the contrasted groups were too small and the overlapping too great to make this conclusion at all certain.

Because of the diversity of treatment and method, it is difficult to summarize in a general way the studies outlined above. It seems fairly certain, however, that investigators are agreed that the pyknic or macrosplanchnic type *tends* toward circular or manic-depressive insanity and that the asthenic or microsplanchnic *tends* toward dementia præcox or schizophrenia. Also it seems probable that the pyknic is more unreserved emotionally than the asthenic.

The Present Experiment

The present experiment is concerned with the three following problems;

(1) The relation of physique (*i.e.* morphologic index and ht/wt ratio) to general intelligence. This is a check on previous work.

[20] A morphologic study of the functional psychoses, *State Hosp. Quart.*, 1924–1925. 10, 413.

[21] The relation of morphology to temperament, *J. Abn. & Soc. Psychol.*, 1924, 19, 254–263.

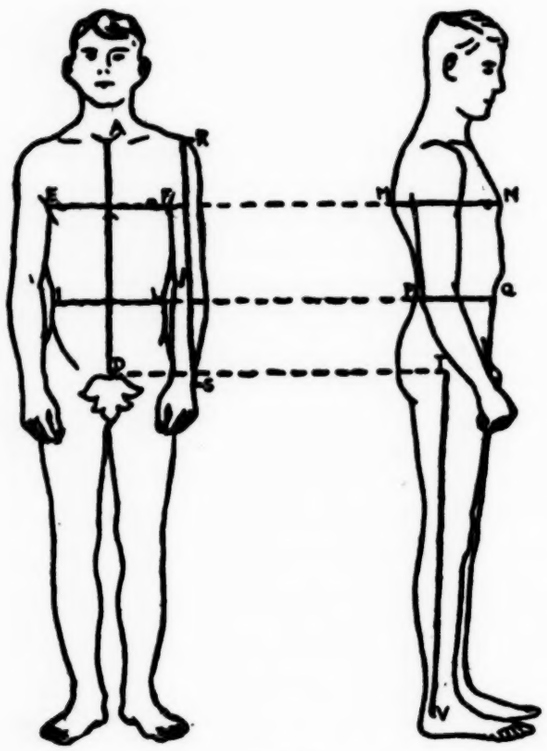
(2) The relation of physique to "emotional instability."
(3) The relation of physique to "social intelligence" or social adaptability.

A total of 221 Columbia College freshmen served as subjects of these experiments. Due to difficulties with the academic schedule, however, it was impossible to secure complete records for each student on all of our tests, and, as a result, the number of cases varies from 122 to 206. The average age of the group was 18.20 years with an SD of 2.00 years. Slightly more than 96 per cent. fell between the limits of 16 and 23 years, while about 99 per cent. fell within the limits 16 to 25 years. Only four men were beyond these limits, the youngest man being 15 and the oldest 29. As a large majority of our subjects fall between puberty and 25 years, it does not appear that the age-factor should materially affect our results.

Calculation of Morphologic Indices.—All measurements used in calculating morphologic indices were taken from photographs of our subjects. These photographs had been made by the gymnasium authorities at the beginning of the academic year with the idea of observing the subsequent physical development of the incoming freshmen. Each picture showed three exposures, a front, a rear, and a side view, the whole print being 12 × 18 cm. The men were, of course, in the nude. All photographs were taken with the subjects in a standing position and were absolutely uniform as regards scale. The camera was solidly bolted to the floor 5 meters from the platform on which the men stood, so that all pictures were taken from the same position and at the same distance from the subject.

Eight measurements were made on each photograph by means of a two-point divider (with adjustable screw) and a millimeter scale. To insure accuracy, the zero point on the scale was notched and the scale fastened to the drawing board. Of the measurements made, seven were used in calculating the morphologic index, and the eighth was a measure of height. The measurements with their points of repère were as follows (see Figure 1):

1. Length of sternum (*AD*): the vertical distance from the suprasternal notch (the upper margin of the sternum) to the upper margin of the pubis (*D*).
2. Transverse chest diameter (*EF*): taken at nipple height.
3. Transverse abdominal diameter (*IL*): taken at the level of the navel.



Measurements taken from photographs (exclusive of height) are indicated by continuous lines.

In order to insure a standard method of measuring, pencil lines were ruled on each print where dotted projections show in illustration.

FIG. 1

ANTHROPOMETRIC BLANK

Name.....	Age.....
PD Score.....	Number.....
Soc. Intell'ce.....	(Gym.) Height (cm).....
Thorndike.....	(Gym.) Weight (kgs).....
	Ht/Wt.....

Measurements Taken

1. Length of trunk (*AD*).....
2. Transverse thoracic diameter (*EF*).....
3. Transverse abdominal diameter (navel) (*IL*).....
4. Antero-posterior thoracic diameter (*MN*).....
5. Antero-posterior abdominal (navel) (*PQ*).....
6. Length of arm (*RS*).....
7. Length of leg (*TV*).....
8. Height (from pictures) (mm).....

Calculations

9. Average 2 and 3 (*EF* and *IL*) = *EL*.....
10. Average 4 and 5 (*MN* and *PQ*) = *MQ*.....
11. Volume of trunk (1 × 9 × 10) = (*AD* × *EL* × *MQ*).....
12. Sum of extremities (6 plus 7) *RS* plus *TV*.....
13. Morphologic Index, $\frac{12}{11}$,.....
14. True Morphologic Index (13 × 32590).....

4. Antero-posterior or sagittal chest diameter (*MN*): taken at the level of the nipples.
5. Antero-posterior or sagittal abdominal diameter (*PQ*): taken at the level of the navel.
6. Length of the arm (*RS*): from the acromion (the marginal point of the shoulder) to the stylion (the wrist joint), with the arm hanging down.
7. Length of leg (*TV*): from the upper margin of the pubis (*T*) to the external malleolus (the ankle).
8. Height: standing height.

Before computing the morphologic index from our photo-measurements, the following preliminary calculations had to be made. A single transverse body diameter (*EL*) was obtained by averaging the transverse chest diameter (*EF*) and the transverse abdominal diameter (*IL*); and a single antero-posterior body diameter, by averaging the antero-posterior chest and the antero-posterior abdominal diameters (*MN* and *PQ*). The product of these two values (*e.g.*, *EL* and *MQ*) was then multiplied by the trunk height (*AD*) to give the volume of the trunk in cubic millimeters. The length of the arm (*RS*) plus the length of the leg (*TV*) gives the length of the extremities; and the "picture" index may be found by dividing the extremities value by the trunk value.

The ratio of the picture-height to the actual body-height of our subjects was found to be 1 : 17.517. This means that 1 cm on the picture scale equals 17.517 cm in actual body dimension: and hence to convert picture measurements into actual body measurements it was necessary to multiply in each case by 17.517. But since the M.I. is a ratio, the numerator being linear and the denominator a cubic expression, in order to convert picture indices into body indices it was necessary simply to multiply by $\frac{17.517}{(17.517)^3}$, *i.e.* by .0032590. To eliminate decimals, this result was then multiplied by 10,000,000 and the first three places taken as the body-index.

Our M.I. differs from that of Naccarati chiefly in that fewer measurements enter into it. One reason for this was the

impossibility of making some of Naccarati's, or better, Viola's measurements on our photographs. To illustrate, in finding the length of the trunk, Naccarati combined three vertical measurements, the length of the sternum, the xipho-epigastric length, and the pubo-epigastric length. It was impossible for us to locate the intermediate points de repère for these measurements, even approximately, and hence we took one measure of trunk-height from terminal points which could be localized. Other cases might be cited.

Even had we been able to take all of Viola's measurements on our pictures, it is doubtful whether the resulting indices would have been better measures of constitution than those we have. Opportunity for error will be seen to loom large in the Naccarati index when it is noted that nine body-measurements are multiplied in sets of three before being added to give the volume of the trunk. In consequence, slight errors in the original measurements will tend to be considerably magnified in the final result, as Naccarati himself realized. He writes (*op. cit.*, 21): "Measurements should be extended to millimeters. An error of 1 mm will reduce or increase the M.I. to about 10 counts. An error of ½ cm plus or minus will cause a change of 50 counts in the index." Elsewhere Naccarati seems to assume that errors in two or three measures will tend to cancel out in the final result—which is valid enough when the measures are averaged, but is certainly not true when they are multiplied together. Wertheimer and Hesketh [22] have criticized the Naccarati index on the score of the possibility of error arising from the large number of measurements taken, and have also questioned the technique employed in making some of the measurements. They observe that "measuring a diameter (the transverse thoracic) at the 'level of the 4th rib' without further instructions becomes difficult when it is considered that the 4th rib curves in three dimensions and varies in position during breathing." [23]

[22] *Op. cit.*, 33 ff.

[23] It is fully realized that our M.I. is open to the criticisms which we have made of the Naccarati index. Errors, of course, are inevitable in any index or combination of measures, since each contributing measure has its own PE. Since our M.I. contains fewer contributing measures than Naccarati's, however, and two multiplications as

The Validity and Reliability of our Morphologic Indices.—(1) The validity of our M.I.s—"picture indices" converted into body indices as described above—was checked as follows.

(*a*) Height was the only actual body measurement which we could secure from the gymnasium records, and hence a direct comparison was first made of these actual body-heights and the converted "picture"-heights. The average actual height of 217 of our subjects was 171.85 cm. The average difference between the gymnasium heights and the converted heights was .809 cm, or about .47 of 1 per cent. of the average body-height. Photographs in which the picture-heights differed from the body-heights (gymnasium) by ± 2.5 cm or more were discarded and not used in any of our computations.

(*b*) The correlation between the unconverted picture-heights (in mm) and the body-heights (gym. records in cm) was .98 ± .002 ($N = 217$).

(*c*) The comparability of our index with that of Naccarati was tested by computing the correlation between our indices and Ht/Wt (both of these measures taken from gym. records) and comparing this r with the r's obtained by Naccarati between his M.I. and Ht/Wt. (In one group of 50 subjects, Naccarati reports this correlation to be .75, in another group of 75 subjects, .70). In our group of 219 men the r between our index of build and Ht/Wt was .81 ± .02. The differences between our r and those obtained by Naccarati are entirely unreliable and might readily arise from fluctuations due to sampling. This suggests that our index is not greatly unlike his based on actual body-measurements.[24] Whether a high correlation between Ht/Wt and M.I. means that the latter is a

against his six, we believe that it is theoretically less open to criticism on this score than his. Again, since our "picture" indices were converted into body M.I.s by multiplying by the constant 32,590, and the first three places in the result taken as the index, the possibility of variations in M.I. due to errors of measurement is considerably reduced. The *average difference* in M.I. between 25 cases measured (and their indices computed) by a different examiner was 5.5 points, *SD* 2.4. A difference of .1 mm in a single measure causes a change of about 2 points in index, which is negligible, probably, in so far as correlations into which these indices enter are concerned. (See further, below.)

[24] The M.I.s in our group ranged from 260 to 654. The average was 407.28, with an SD of 53.88. Naccarati reports a spread of 435 to 721, and of 360 to 671 in his two groups of 75 and 50.

good estimate of constitution depends, of course, upon what one wishes to measure. Naccarati was inclined to stress the fact that Ht/Wt is an "approximate indicator" of morphologic index, and he uses his r of .228 between Thorndike score and Ht/Wt as an argument in favor of a relation between general intelligence and physique.

(2) The reliability of our indices was determined as follows. A sample of 25 cases, randomly selected, was remeasured by a different examiner and the M.I.s computed in entire independence of the first set of measurements. The correlation between these two sets of M.I.s was .96 ± .01. With the aid of Kelley's formula [25] we estimate the reliability coefficient in our group of 219 to be .97 ± .01.

Description of the Tests Used.—(1) The Thorndike Intelligence Examination for High School Graduates. This is the standard intelligence examination in use at Columbia for the past seven years as an entrance test for freshmen. It consists of three parts (generally) and requires about 3½ hours to administer. This examination has been very fully described by Wood.[26]

Thorndike scores were available in the files of the Admissions office for 206 of our 221 men. The average score for the group was 82.09, with an SD of 12.14. Scores ranged from 47 to 116.

(2) The Woodworth Personal Data Sheet or Psychoneurotic Inventory. This test consists of 116 questions dealing with delusions, obsessions, fears, anxieties, etc. Each question is to be answered simply as Yes or No. The number of questions answered "wrongly" (with neurotic implication) by the average normal person is around 10 or 11, by the psychoneurotic and the insane, 30 or more.[27]

This test was given to 151 freshmen. Scores ranged from 0 to 43, with an average of 13.82 and an SD of 9.20.

(3) The George Washington Social Intelligence Test. This test, which is designed to gauge one's social perception and knowledge of social usages and amenities, consists of six parts:

[25] T. L. Kelley, *Statistical method*, New York, 222.

[26] *Measurement in higher education*, New York, 1923.

[27] S. I. Franz, *Handbook of mental examination methods*, New York, 1919, 170.

1. Judgment in social situations
2. Memory for names and faces
3. Recognition of mental states from facial expression
4. Observation of human behavior
5. Social information, and
6. Recognition of the mental state of the speaker.

About 40 minutes are required to give this test. It was administered by us to 123 freshmen, whose scores ranged from 66 to 132 (the maximal score is 160). The average score was 105.04, and the SD 14.34.

Reliability of the Tests. The reliability coefficients of our three tests were as follows:

	Thorndike	P.D.	Social Intell'ce
Reliab'y Coeff.	.84	.90	.79
SDs	12.14	9.20	14.34
N	206	151	123

Results.—The chief results of this study are expressed in the form of correlations between the various measures (Table I).

TABLE I

INTERCORRELATIONS OF THE MEASURES

Tests	*N*	*r*	*r* (corrected for attenuation)
M.I. & Thorndike Score	206	.074 ± .047	.082
M.I. & Social Intell.	123	− .055 ± .061	− .063
M.I. & P.D. Sheet	151	.047 ± .060	.050
Ht/Wt & Thorndike Score	204	.101 ± .047	.110
Ht/Wt & Social Intell.	122	.051 ± .061	.057
Ht/Wt & P.D. Sheet	150	.088 ± .052	.093
Ht/Wt & M.I.	219	.810 ± .016	.822
Thorndike score & Social Intell.	118	.420 ± .051	.515

The highest *r* in the table (disregarding the last two) is that between Ht/Wt ratio and the Thorndike score. The corresponding *r* between M.I. and Thorndike score is slightly lower. These *r*'s are closer to the correlations between physique and general intelligence obtained by Heidbreder and Sheldon (see above) than to those got by Wood, Naccarati, or Mohr and Gundlach. All of the correlations are negligible in size.

Not only are our measures of constitution negligibly

related to general intelligence, but they are related to emotional instability and social intelligence to no greater extent. The correlation of .42 between Thorndike score and Social Intelligence is probably to be expected in view of the fact that the latter test, in spite of the ingenious selection of its component tests, is certainly a measure of "abstract" or verbal intelligence as well as of 'social' intelligence.

In casting about for a possible explanation of the low correlations in Table I the factor of age first suggests itself. Age can hardly be an important factor, however, as 90 per cent. of our group were between 16 and 20 years old. Besides, ruling out the influence of age through partial correlation, if it did anything at all to our obtained r's, would probably serve simply to reduce them still further. Nor does it appear that inaccuracies in the M.I.s can be the effective cause of the low correlations for two reasons: (1) in our correlation diagrams the M.I.s were tabulated in steps of 20 units, so that considerable error might be present in a single measure without removing it from its proper step-interval; and (2) the correlations of Ht/Wt with the other variables are also very low, although both height and weight were taken from the gymnasium records.

The possibility of non-linear regression remains to be considered. Curvilinearity of regression has been found to obtain to a greater or lesser degree in other studies of mental and physical traits.[28] To settle this question for our data, we computed correlation-ratios for each pair of variables as shown in Table II.

Table II

Correlation-ratios between the Different Measures

Variables					ηxy
Y	X	N	r	ηxy	(corrected)
M.I. and Thorndike Score		206	.074	.198	.090
M.I. and P.D. Sheet		151	.047	.237	.171
M.I. and Social Intelligence		123	−.055	.212	.124
Ht/Wt and Thorndike Score		204	.101	.273	.118
Ht/Wt and P.D. Sheet		150	.088	.186	.095
Ht/Wt and Social Intelligence		122	.051	.346	.193

These correlation-ratios are surprisingly high, even when

[28] Sommerville, *op. cit.*, 75–81.

corrected for fineness of grouping (*e.g.*, by the $\frac{\text{zeta}}{\text{PE(zeta)}}$ test); [29] but none of them is sufficiently larger than its corresponding r to denote a difference which could not arise from sampling alone. Several of the uncorrected η's, however, (*e.g.*, those which involve the ratio Ht/Wt) pass the Blakeman test for non-linearity of regression. But as this test is not generally applicable unless $\eta^2 - r^2$ is small as compared with r (which is certainly not true for our table), it is extremely doubtful whether this rougher test may be used legitimately.[30] In short, we are forced to the conclusion that there is no real evidence of significant relationship—linear or otherwise—between our indices of build and our tests of general intelligence, social intelligence, and emotional instability.

One further check was made on this general result. Table II shows that the highest calculated η is that between Ht/Wt and Social Intelligence; and this coefficient is the least reduced by the fineness of grouping correction. We decided to test the non-linearity of regression between these two variables by dividing the 122 men, for whom we had Social Intelligence ratings, into three groups on the basis of their Ht/Wt ratios, and computing the average social intelligence score for each of these groups. The upper 25 per cent., who constituted the first group, are called "microsplanchnics," the middle 50 per cent. "normosplanchnics," and the lower 25 per cent. "macrosplanchnics." It is extremely doubtful whether these terms apply here strictly, as our group is small and highly selected, and it may be that all of its members are either microsplanchnic or micro- and normosplanchnic. But it is probably fair to consider the upper 25 per cent. to be relatively microsplanchnic as compared with the lower 25 per cent.; and hence these terms are not misleading if understood in a relative rather than an absolute sense.

The following limits in Ht/Wt ratio were obtained for the three groups. The average scores on the Social Intelligence

[29] Kelley, *Statistical method*, 238.

[30] H. L. Reitz, *et al.*, *Handbook of mathematical statistics*, New York, 1924, 131.

	Upper 25 per cent. (micro)	Middle 50 per cent. (normo)	Lower 25 per cent. (macro)
Ht/Wt Ratio Limits	3.92–3.01	3.00–2.60	2.59–1.88

Test made by the three groups, the SDs of these distributions and averages, and the reliability of the differences between the averages, are given in the following table.

Table III

Comparison of the Social Intelligence Scores made by the Three Physique Groups

Classified in terms of Ht/Wt	*N*	Social Intell.	SD(dis)	SD(aver)
Microsplanchnic (upper 25 per cent.)	31	103.63	12.87	2.314
Normospl. (middle 50 per cent.)	61	107.34	15.26	1.952
Macrosplanchnic (lower 25 per cent.)	30	103.00	12.41	2.266

Reliability of the Differences

	Obtained Difference	SD (diff)	D/SD (diff)
Normo—macro	4.34	2.99	1.45
Normo—micro	3.71	3.02	1.22
Micro—macro	.63	3.23	.19

Table III indicates a slight—and statistically unreliable—superiority of the "normosplanchnic" group over both the "microsplanchnic" and the "normosplanchnic" groups. It is interesting to speculate that this difference is due to the relatively better physiques of the middle group which would lead to greater ease and adaptability in social situations. But such a conclusion is entirely hypothetical, and is not supported statistically. The existence of non-linear relationship between Ht/Wt and Social Intelligence Test is suggested, to be sure, by the increase and then decrease in average score as we go from the high to the low index groups. While it cannot be shown for our group, it is possible that in a very large group, randomly selected from the general population, curvilinearity of regression might be established.

Summary

1. Anthropometric measurements were taken from the photographs of 221 Columbia College freshmen and used to calculate a "morphologic index" based upon that of Nac-

carati. These indices together with the Ht/Wt ratios of the same group were then correlated with the Thorndike Intelligence Examination scores made by 206, with the Social Intelligence Test scores made by 123, and with the Woodworth Personal Data Sheet scores made by 151 members of the group.

2. Small and unreliable correlations ranging from .051 and .101 (for Ht/Wt) and from − .055 and .074 (for M.I.) were obtained between Ht/Wt, M.I. and the other three variables.

3. Correlation-ratios calculated from our tables, while considerably larger than the product-moment *r*'s, were not large enough to pass the tests for establishing non-linearity of regression. It is suggested that the correlation between Ht/Wt and Social Intelligence Test might be curvilinear if calculated from a larger and less highly selected group.

4. A correlation of .42 (.52 when corrected for attenuation) was obtained between Thorndike scores and the Social Intelligence Test. This suggests that the latter test is, to a certain degree, a measure of intelligence.

5. The correlation between our M.I. and the Ht/Wt ratio was .81, indicating that the two measures of physique measure virtually the same thing.

(Manuscript received September 10, 1927)

GALVANIC PHENOMENA OF THE SKIN

BY LLOYD A. JEFFRESS

University of Texas [1]

From the earliest time in the history of the 'psycho-galvanic reflex' (as it was long called) there has existed a doubt as to whether the phenomenon discovered by Féré (4) and exhibited when a galvanometer is connected in circuit with the body and with an electric battery, and that first observed by Tarchanoff (10) and obtainable when the body alone is connected in circuit with the galvanometer, are two aspects of one and the same phenomenon, are two distinct phenomena or are causally related phenomena. Concensus of opinion seems to point toward the last conjecture, but with one exception [2] no experiments have been undertaken with a view to measuring directly this relationship.

The present series of experiments was undertaken to determine to what extent the two phenomena are related. Three general problems are presented; (1) Is there any relation between the size of the galvanometric deflection with and without an external source of electromotive force? (2) Are the latent periods of the two phenomena the same? and (3) Can an hypothesis be advanced, or has any hypothesis been advanced, which will account for all of the facts?

In order to approach the first problem it is necessary to

[1] The work was done at the Psychological Laboratory of the University of California. The writer desires to express his gratitude to Professor Warner Brown for his assistance and counsel throughout the course of these experiments, to Professor S. S. Maxwell for many suggestions and for his permission to use the electrocardiograph, and to Dr. Donald A. Macfarlane for his coöperation both as subject and co-worker.

[2] The exception is an experiment mentioned by Farmer and Chambers (3) in which they measure the size of the deflections obtained from the two hands simultaneously with two instruments, one connected to the hand directly and one through a Wheatstone bridge. They found a high degree of relationship between the two, but make no mention of the actual size of the correlation. The present series of experiments was undertaken before the appearance of this article, but not before it was submitted for publication.

devise some method for recording the two phenomena at the same time. The following apparatus was used; a Hindle electrocardiograph with the usual control box,[3] a Leeds Northrup D'Arsonval galvanometer (Model 2500E) with a sensitivity of 631 megohms and a resistance of 560 ohms, a camera supplied with the electrocardiograph to record the excursions of the galvanometers, necessary resistances to make up a Wheatstone bridge, a shunt and various accessories.

The apparatus was so arranged that the shadow cast by the string of the electrocardiograph and the beam of light reflected by the mirror in the D'Arsonval galvanometer fell on the scale and slit of the camera and registered on the film as a light and a dark line, respectively, on a gray background.

To make it possible to record the exact moment of stimulation, a signal magnet was provided, so arranged as to cast a shadow across part of the slit, and so connected that it was automatically actuated when the stimulus was given. A time marker measured the time in 1/25th sec., with an estimated 1/100 sec. When less accurate measurement of time was desired a Jaquet chronograph was substituted and the film run at a slower speed (5 mm a sec.).

The experimenter could connect either galvanometer or electrocardiograph to either of two pairs of electrodes in such a fashion that one pair was in connection with a Wheatstone bridge and the other directly with a galvanometer. The arrangement usually employed was that of using the D'Arsonval with the bridge while the Hindle was directly connected to the subject. This was done because no potentiometer was at hand to use with the D'Arsonval galvanometer in order to compensate for body-potential; such a potentiometer is included in the control box of the Hindle or string galvanometer. An additional reason for using the string galvanometer in measurements of potential is that the changes in potential are apt to be more rapid and are not so satisfactorily indicated on the D'Arsonval galvanometer as are the more gradual changes of resistance.

Switch *B* in the figure is used to connect either pair of electrodes to the galvanometer and the other to the bridge. Switch *A* connects either galvanometer to the bridge and the other directly to the electrodes. The commutator shown at the bottom of the figure reverses the current through the subject without changing the direction of its flow through the bridge or galvanometer.

In addition to the two pairs of electrodes, there is a fifth (labelled "mouth" in the figure) which is placed in contact with the mouth and which may be used in conjunction with any of the other four electrodes and the string galvanometer by setting the control switch *C* and switches *A* and *B* at the appropriate positions.

The Wheatstone bridge used was of the usual form. To avoid changing the size of the current through the body, any adjustments which were necessary were made by changing the resistance placed in series with the subject in the fourth arm of the bridge, according to the method suggested by Prideaux (8). The bridge was balanced at more than the necessary resistance, (usually 15,000 ohms) and the difference between this value and the resistance of the subject was added by means of a series resistance box.

[3] A modification of the string galvanometer invented by Einthoven, built by the Cambridge & Paul Instrument Company from specifications drawn up by Hindle. The resistance of the string used in these experiments was 2210 ohms.

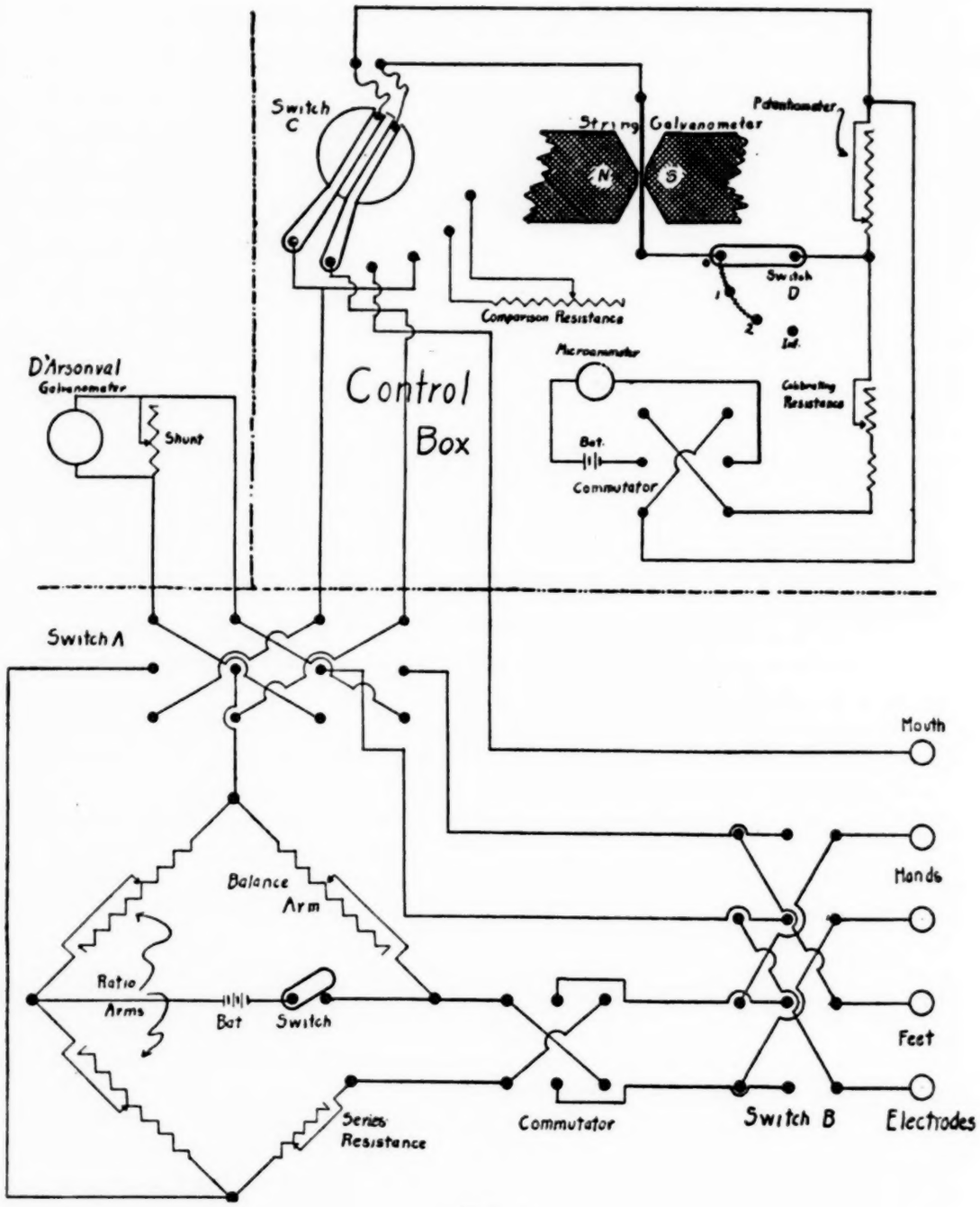


FIG. 1

In practice, since the scale of the camera was short (6 cm) in proportion to its distance from the D'Arsonval galvanometer (approx. 1 meter), it was found possible to use the size of the deflection of the galvanometer as a measure of the size of the change of resistance. Before starting a series of experiments the shunt across the galvanometer was set at such a value that a given deflection indicated a given change of resistance. This relation did not change during the experiment since the balance of the bridge was maintained by compensating, through the series resistance, the changes in

the subject's resistance. Thus by comparing the graphic record with the record of the resistance box settings it was possible to tell the resistance of the subject at any time during the experiment. It is not necessary to know the value of the shunt resistance, since it is adjusted empirically until the desired sensitivity of the galvanometer is reached. This fact makes it possible to employ as a shunt a high resistance rheostat instead of a calibrated resistance box.

In calibrating the galvanometer and bridge it is necessary to substitute for the subject a metallic connection and balance the bridge with this in the circuit; otherwise, when a change is made in the resistance of the subject's arm of the bridge, there is a consequent change in the current flowing, and, accordingly, of the counter-electromotive force of polarization, a change which causes the calibration to be either too low or too high according to whether the calibrating change was in the form of a decrease or an increase in the resistance in the subject's arm of the bridge.

An additional reason for observing the above precaution is that in making the calibration it is necessary that the resistance in the circuit should remain constant except for the change desired. This is not possible when the subject is connected to the apparatus, since there is always apt to be a slight variation due to unnoticed stimuli.

The control box used in conjunction with the string galvanometer is shown (in the Figure) surrounded by dotted lines. It consists of the switch *C* which is used to connect any pair of the three lead wires to the galvanometer; the comparison resistance, (which is a resistance box that may be substituted, by turning *C* to the extreme right for the external circuit and thus used to measure the resistance of the external circuit by the method of substitution), a potentiometer capable of giving potentials by .0001-volt steps to .1 volts, and a switch *D*, which is used to connect the string to the circuit through two steps of protective resistance. The potentiometer circuit consists of a battery, a commutator for reversing the current, a calibrating rheostat and microammeter for adjusting the current through the potentiometer to the correct value, and the potentiometer. As can be seen from the diagram, when the potentiometer is set for any value other than zero it introduces an additional resistance in the circuit; but since the resistance of the calibrating rheostat is high in proportion to this the error introduced is negligible.

By adjusting the tension on the string it is possible to obtain any degree of sensitivity from about 1 mm per millivolt across the string terminals to about 50 mm for 0.1 millivolt. The value usually employed was the one which is standard in electrocardiographic practice, namely, one centimeter per millivolt with the subject in the circuit, which means, in the usual terminology, one scale division to about 1×10^{-7} ampères, assuming the resistance of the subject to be 8,000 ohms. This sensitivity is sufficient for most of the work on the potential effects of the body, though when the apparatus was used in connection with the Wheatstone bridge a more sensitive adjustment was necessary.

The electrodes employed were of the Ostwald type and were constructed as follows. A liter beaker was filled to a depth of about a centimeter with clean dry mercury. Into this was thrust a glass tube with a platinum wire sealed into the end to make contact with the mercury. The other end of the platinum wire was connected to the galvanometer by means of a copper wire which dipped into a little mercury in the tube above the platinum. Above the mercury in the beaker was a layer of precipitated mercurous chloride, and above this was the solution, into which the subject dipped his hand, of 20 per cent. sodium chloride. It was found that great care is necessary to prevent the junction of the mercury and platinum from becoming moistened by the solution and thus acting as an electric cell and setting up a large potential.

To obtain mercurous chloride which is free from the impurities usually found in commercial calomel it was prepared in the laboratory. Mercury was dissolved in C.P. nitric acid leaving an excess of mercury. To this mixture was then added hydrochloric acid in sufficient quantity to combine with all of the mercurous nitrate, still leaving an excess of mercury to avoid forming the mercuric chloride. After the precipitate of mercurous chloride had settled, the liquid above it was syphoned off and the mercury and calomel were washed repeatedly in sodium chloride solution of the strength later to be used in the electrodes. The calomel and the salt solution were then agitated together to get the precipitate in suspension for the time being and the mixture was poured on top of the mercury in the electrodes. The calomel settled and formed a layer above the mercury.

The electrodes were then tested by connecting the platinum wires of two of them to the two poles of the string galvanometer. If the potential difference indicated when the two solutions were connected by moist filter paper was larger than about one millivolt, the electrode responsible was discarded and a new one built until all five gave values within this range when connected in the necessary permutations. Inequalities in the concentrations of the solutions in the various electrodes were avoided by periodically connecting the solutions with each other by tubes filled with the liquid, and letting them stand that way when not in use.

The subject was seated in a chair to the arms of which were attached the electrodes for the hands. His hands lay along the arms of the chair with the fingers dipping into the electrodes. The mouth electrode was swung on an arm from the wall and was adjustable as to position so that it could be brought into close juxtaposition with the mouth and connected to it by a piece of moistened filter paper. Care was taken to keep the area of the hand immersed in the electrode constant by using rubber gloves with holes for the portions of the fingers to be immersed. It was found that the subject could maintain the required posture for ten or fifteen minutes without moving appreciably.

A first series of experiments was undertaken to ascertain what effect movements and other irrelevant activity on the part of the subject have upon the galvanometer. It was found that with few exceptions there was little chance of the deflection resulting from a movement being confused with the 'psychogalvanic' reaction. Gross bodily movement had practically no effect unless accompanied by a conspicuously strong tension of the muscles in the arms. Such movements were reported by the subject and the reaction discarded. Usually, too, such a movement resulted in a curve of irregular shape which was quite different from the characteristic curve of the psychogalvanic reflex.

The effect of taking a deep breath or of sighing is indistinguishable from the true reflex in both form and latent period and is thought by many observers, among them Wechsler, to be of the same general nature.

One type of muscular activity found to have a pronounced galvanic effect is clenching the hand. This results in the hand clenched becoming positive to the rest of the body. The phenomenon was first observed by Du Bois Reymond (2). It is in the direction opposite from the current set up by the Tarchanoff reaction and hence is easily distinguished from it.

Nature of the Potential Deflection

During some of the preliminary experiments symmetrical parts of the body were used and it was found that the deflections obtained by Tachanoff's method under these circumstances are highly irregular and unpredictable and very different from those obtained when the subject was connected to the bridge. The latter were uniform in shape and direction and always in the direction of decrease of resistance.

When the two hands were connected to the string galvanometer and the difference of potential between the hands was reduced to zero (by opposing it with an equal potential in the opposite direction) and then a stimulus was given, the direction, shape and size of the resulting deflection seemed to be governed by chance. These deflections could not be made more regular or predictable by changing the relative area of immersion of the two hands.

These results are not inexplicable if we conceive of the potential as being set up at the skin, and regard the skin as an electric cell. In the cell made, say, of copper and zinc the potential produced is not a function of the areas of the metals immersed but simply of the chemical composition. If an analogous condition exists in the skin and the potential produced is approximately the same at symmetrical parts of the body, the difference of potential between those parts will be small. If, further, the potential set up is variable and varies in roughly, but not exactly, the same manner in symmetrical parts, the size and direction of the potential difference between these parts will vary in an irregular fashion. Such a small and irregular difference of potential was found.

We may assume further that symmetrical parts of the body

are more like each other in potential and variation of potential than are asymmetrical parts. From this we may infer that by employing asymmetrical parts of the body, one of which has a low, and the other a high potential, we should obtain deflections which are predominantly in one direction.

Since it is very likely, as the literature shows, that the Tarchanoff phenomenon is due to the activity of the sweat glands, we should expect the regions rich in these glands to exhibit higher potentials than those poor in this provision. Thus by placing one electrode in contact with some part of the body rich in sweat glands and the other in a poor region we might obtain results which would throw light upon the nature of the potential effect.

The region chosen for the neutral electrode was the mouth. An electrode identical in construction with those used for the hands was employed and placed in contact with the mouth by means of a strip of filter paper moistened in the electrolyte. This was laid on the tongue and care was taken to prevent its coming in contact with the skin. Some trouble was experienced by the subject in keeping the contact between the paper and the mouth constant, but, by practice, one of the subjects was able to remain very quiet and to withstand the severest stimuli without movements large enough to become apparent upon the galvanometer. The hands were placed in electrodes, either of which could be connected to the galvanometer in opposition to the mouth or other hand.

The difference of potential between the mouth and either hand was found to be much greater than that between the two hands; the initial value was frequently as high as ten or fifteen millivolts as contrasted with two or three between the hands and the difference of potential was always in the same direction; the hand was always negative to the mouth, that is, the current set up was centripetal in the hand.

The result of stimulation was, with a few exceptions, in the direction of an increase in this centripetal current; that is, the hand became more negative.[4]

[4] These results are in accord with those of Wells and Forbes (15).

Relation between Deflection With and Without the Battery

We are now in a position to attack the question of the relation between the size of the deflections obtained with and without an ectosomatic source of electromotive force.

Both galvanometers were employed. The D'Arsonval was connected to the Wheatstone bridge, and the string directly to the subject. The latter was connected between the mouth and one hand, and the bridge was connected to two electrodes which were in contact with fingers of the other (right) hand.

Break shocks from the inductorium were used as stimuli. The inductorium was so arranged that the size of the shock could be varied in order to assure a variety in size of response. The wires from the secondary of the inductorium were connected to two copper buttons which were placed in contact with the subject's forehead. The shock afforded was distinctly unpleasant.

The relation between the size of the potential deflections and the size of the resistance deflections was very close. When one of the deflections was large the other was proportionally large and when one was small the other was small. When, at the end of the series, the subject was adapted to the stimulus both deflections were too small to measure. To determine the degree of relationship, a coefficient of correlation was computed for a series of twenty-two reactions and found to be .958 ± .012.

In another experiment the two electrodes from the string galvanometer were connected to two parts of the same hand; the fingers and the thumb. The remainder of the apparatus was arranged as just described. Three series of similar reactions were obtained.

The difference of potential between the fingers and thumb was much smaller than it had been between the hand and mouth; but there was sufficient dissimilarity between the parts to permit clear reading of the reactions. The potential changes were found to correlate very highly with the resistance changes as they did in the previous experiment. The correlations with various subjects were .959 ± .014, .911 ± .015 and .981 ± .004. There is considerable variation in the absolute size of the reactions among subjects, and furthermore, the one type of reaction may be relatively weaker, as compared to the other type, in one subject than another; but the

relation between the magnitudes of the two types remains very high for a given subject.

Latent Periods

Due to the fact that both galvanometers could not be connected to the same parts of the body at the same time, the following method was employed in the study of latent periods. The string galvanometer was used in all cases, and connected part of the time directly to the subject and part of the time to the Wheatstone bridge.

The string was chosen because of its short period. The stimulus used was the electric shock. The apparatus was so arranged that only the break shock from the inductorium was used, and so that breaking the circuit through the inductorium simultaneously broke the circuit through the signal magnet. In this way the moment of stimulation was accurately recorded.

The film in the camera was run at a high rate of speed so that one mm corresponded roughly to 1/25 sec. Measurements of latent period were made to 1/25 sec; the comparatively slow movement of the string at the beginning of a reaction making it impossible to estimate any more accurately the point of beginning of the deflections.

No difference was found between the two sets of measurements when taken at the same part of the body. It was found, however, that measurements taken with the electrodes connected to the feet and to the hands differed. The former averaged for one subject, 1.95 sec. and the latter 3.10 sec. Similar differences were found in several subjects. No explanation is at hand.

In order to be certain that the method used above was free from the possible error that the latent period might vary over a long series of reactions, the following experiment was performed. The latent period of each of the reactions in a long series was measured and the means obtained of the first and the last half of the series. The results were as follows:

Mean of first half of series	1.808 sec.
Mean of second half of series	1.816 sec.
Difference	0.008 sec.
Standard error of difference	0.067 sec.

It seems clear that no measurable error results from comparing one type of reaction as measured at the beginning and the other type as measured at the end of a series.

Interpretation of Results

It has been shown by Aebly (1), Gildemeister (5) and Weiss (14) that Féré's phenomenon depends upon an alteration in the counter electromotive force of polarization,[5] an alteration which is probably the result of activity of the sweat glands.

It has been suggested further that Tarchanoff's phenomenon depends also upon the activity of the sweat glands. Some evidence [6] has been shown for this hypothesis, and our experiments tend to support it.

[5] In this connection, Prideaux (8) attempts to distinguish between counter electromotive force of polarization and resistance of polarization, and quotes Waller (12) as follows;

"If a potentiometer is put in series with the subject in the fourth arm of the bridge, we shall obtain certain indications that will enable us to decide between the two alternatives—electromotive force increased or resistance decreased. If the former, a known E.M.F. from the potentiometer should give the same deflection before and during an emotive change. If the latter, the potentiometer deflection should be increased during the emotive change. And in point of fact this second alternative is what actually occurs; whenever I have measured the resistance in ohms before and during emotive excitement, I have found that the potentiometer deflection was inversely proportional with the resistance." Waller, in making this statement, and Prideau, in accepting it, have naïvely assumed that a Wheatstone bridge behaves in the same fashion as a simple circuit. If we were dealing with a simple circuit Waller's statement as to what should be expected in the case of a change in resistance as against a change in potential would obtain; but in the case of a bridge it does not hold. From a consideration of Kirchoff's laws it is apparent that any change in either the resistance or the potential in any arm of the bridge changes the whole set of relations among the currents in the other arms of the bridge, and through the galvanometer. Thus it is impossible to assume any such simple state of affairs as Waller postulates in the above statement. The actual calculations of the deflections to be expected is laborious and depends upon the knowledge of all of the resistances and potentials in the bridge except the one value to be solved for. It is impossible to say in Waller's case just what the values were since he gives in his article none of the figures.

This experiment has been considered at considerable length since it is often quoted as proving that the galvanic reflex does not depend on a change of counter electromotive force of polarization. When this notion is combined with the fact, already demonstrated, that the galvanic reflex does depend upon polarization we are led to the conclusion that the kind of polarization which causes the galvanic reactions can occur without setting up a counter electromotive force. This seems to have been the point of view of Prideaux; but it is entirely without foundation in physical fact.

[6] The recent work of Darrow (1*a*), Sensory, secretory and electrical changes in the skin following bodily excitation, this *Journal*, 1927, *10*, 197–226, shows that there is a fairly high correlation between the amount of moisture (sweat) present at the surface of the skin and the galvanic reaction as measured without an external source of E.M.F.

The writer believes that the evidence in favor of the sweat gland theory for Féré's phenomenon is conclusive, and that in view of the high correlation between the two phenomena the theory must also be evoked to account for Tarchanoff's. It remains to consider the possible physical and physiological mechanisms by which the reactions are effected.

Wechsler (13) proposes a theory to account for Tarchanoff's phenomenon in terms of change of polarization, maintaining that Tarchanoff's is but a special case of Féré's phenomenon. He says, in effect, that the mere presence of a galvanometric deflection in the absence of an ectosomatic source of electromotive force is inadequate proof that the phenomenon is actually due to a new potential, that when any two parts of the body are connected to a galvanometer an electromotive force is set up, and that it makes no difference whether the current is endosomatic or ectosomatic in origin so far as the counter effects of polarization are concerned. He then goes on to say that alterations of *this* counter E.M.F. can take place just as they do when an external source of current is used. Wechsler's exact words are "If the galvanic response obtained without exsomatic (!) current is identical with that obtained when an external source of current is introduced, namely, is of a nature of a diminution of the counter E.M.F. set up against it, then we should expect that the direction of the galvanic response maintained a constant relation with the respect to the direction of the galvanometric deflection produced by the initial body current."

In other words, whatever deflection occurs when the subject is stimulated, should be, according to Wechsler's theory, in the same direction from zero as the initial deflection which was due to the difference of potential between the two parts of the body. When, therefore, the subject is connected to the galvanometer and there is a deflection to the right, upon stimulation the galvanometer should be deflected still more to the right.

Wechsler found that in most cases the deflections were as he had predicted. The other cases and the observations of

others, as Radecki (9), which were at variance, he puts down as not being 'psycho-galvanic' reactions.

That Wechsler's theory is not tenable can be seen from the fact that the Tarchanoff phenomenon occurs when there is no polarization, because no current is flowing. This condition was established by neutralizing the body potential by an external potential obtained from the potentiometer. Furthermore, the phenomenon continues in the same direction when the current flowing (and hence the counter E.M.F. of polarization), is in the opposite direction. Under the latter conditions, according to Wechsler's theory, the galvanic deflection should be reversed.

Three general hypotheses have been offered to account for Féré's phenomenon, in connection with the activity of the sweat glands. (1) That of Peterson and Jung that the increase of the secretion of sweat serves to decrease the resistance of the skin by the production of a better path in the form of sweat for the flow of current. Thus, according to their theory, we should expect a simple decrease in resistance. As we have seen, this is not the case, since the phenomenon occurs without a change in resistance, but rather a decrease in the counter electromotive force of polarization. (2) That of Waller (11), that there is an increase in the size of the pores, and hence a decrease of resistance. This is open to the same objection as the above, since it takes no account of polarization. (3) That of the semi-permeable membrane. Several experimenters have mentioned the semi-permeable membrane, but the writer has seen no account in which it has been connected with polarization and with the Tarchanoff phenomenon exhibited in the absence of an ectosomatic potential.

The concept of the semi-permeable membrane seems a useful one to account for the phenomena observed. If we consider the interior membrane of the sweat glands as semi-permeable, we may describe the phenomenon of secretion as an alteration of permeability which permits the passage of the perspiration from the interior of the sweat glands into the tubules. If we further assume that the membrane is perme-

able to only one of the ions in the sweat, which is already in the tubules, or in the electrolyte of the electrodes, we may explain the phenomenon of polarization which takes place when a current is caused to flow through the skin. If the membrane is permeable to only one of the ions, say the cation, the others collect at the membrane and oppose the flow of current; in other words, the membrane becomes polarized, and a counter electromotive force is set up opposing the flow of current. Now if upon stimulation the permeability of the membrane is changed so that it is more or less permeable to both ions, the current would flow with less difficulty; *i.e.* the counter electromotive force would be reduced. This, as we have seen, is what actually occurs upon stimulation.

That secretory activity of the sweat glands is brought about by nervous activity is evident, and that the action of nerves is accompanied by electrical effects is well known, so that it seems reasonable to suppose that the action of the nerves supplying the sweat glands is an electrical action. We should expect, then, that stimulation of the sweat glands would exhibit itself by an electrical manifestation or "action current" of the gland itself. This has been shown to be the case. Several experimenters, Waller (11) and Mendelsohn (6) have shown that activity of the sweat glands is accompanied by an increase in electrical potential directed from the surface of the skin inward. This has been shown in the case of the cat connecting the paw and an indifferent part of the body to the galvanometer and stimulating the sciatic nerve, which innervated the sweat glands in the pad of the foot. When this is done, the foot becomes negative to the indifferent part of the body.

These results are in complete agreement with our own. We found that the hand, for example, becomes more strongly negative to the indifferent part of the body (mouth) when the subject is stimulated, thus leading us to believe that we are dealing here, as in the case of the cat, with action currents of the sweat glands.

We may therefore account for the galvanic phenomena in question as follows. The reactions obtained when no external

source of potential is used (Tarchanoff's phenomenon) are action currents of the sweat glands. The activity is shown electrically by a current directed inward from the surface of the skin. It is accompanied by a change in permeability of the membrane of the glands which in turn results in a change in electrical polarizability. This latter exhibits itself electrically as a variation in counter electromotive force with the consequent variation in apparent resistance which is known as Féré's phenomenon.

This hypothesis accounts for the presence of the two phenomena and for the high degree of relation between the two both as to magnitude and as to other characteristics as well.

Summary

1. There is a high degree of relationship between the magnitude of the deflections obtained with and without an ectosomatic source of potential. The coefficient of correlation indicating the size of this relationship is of the order of .95.

2. The two phenomena exhibit the same latent period.

3. An hypothesis has been suggested which accounts for the close relation between the phenomena, as well as for discrepant results obtained by a number of other investigators.

References

1. Aebly, J. Zur Analyse der physicalischen Vorbedingungen des psychogalvanischen Reflexes mit exosomatischer Stromquelle. (Inaug. Diss.) Zurich, 1910.

1*a*. Darrow, C. W. Sensory, secretory, and electrical changes in the skin following bodily excitation, *J. Exper. Psychol.*, 1927, *10*, 197–226.

2. Du Bois Reymond. *Untersuchungen ueber thierische Elektricität.* Berlin, 1849, 261–276.

3. Farmer, E., & Chambers, E. G. Psychogalvanic reflex in psychological experiments, *Brit. J. Psychol.*, 1925, *15*, 237.

4. Féré, C. Note sur les modifications de la résistance électrique sous l'influence des excitations sensorielles et des émotions, *C. r. Soc. de Biol.*, 1888, *40*, 217.

5. Gildemeister, M. Ueber elektrischen Widerstand, Kapazität und Polarization der Haut, *Arch. f. d. ges. Physiol.*, 1909, 176, 84–105.

6. Mendelsohn, —. Richet's *Dictionnaire de physiologie*, Article on *Electricité*, Vol. 5, 350.

7. Peterson, F. & Jung, C. G. Psychophysical investigations with the galvanometer, *Brain*, 1907, *30*, 153–218.

8. Prideaux, E. The psychogalvanic reflex; a review. *Brain*, 1920, *43*, 50–73.

9. Radecki, W. Recherches expérimentales sur les phénomènes psychoélectriques, *Arch. de Physiol.*, 1911, *11*, 209–295.

10. TARCHANOFF, J. Ueber die galvanischen Erscheinungen an der Haut des Menchen bei Reizung der Sinnesorgane und bei verscheidenen Formen der psychischen Tätigkeit, *Arch. f. d. ges. Physiol.*, 1890, *44*, 46–56.
11. WALLER, A. D. The influence of drugs upon the electrical conductivity of the palm of the human hand, *Proc. Roy. Soc.*, 1919, *91*, 32–40.
12. WALLER, A. D. *Signs of Life.* N. Y., 1903, 97.
13. WECHSLER, D. The measurement of emotional reactions, *Arch. of Psychol.*, 1926, no. 76.
14. WEISS, —. Cited by Guilleminot in *Electricity in medicine.* London, 1906, 202.
15. WELLS, F. L., & FORBES, A. On certain electrical processes in the human body and their relation to emotional reactions, *Arch. of Psychol.*, 1911, no. 16.

(Manuscript received August 1, 1927)

A STUDY OF THE LEARNING CURVES FOR TWO SYSTEMS OF SHORTHAND

BY PARRISH LITTLE JETTE

Columbia University

The course of acquisition in typewriting has long been a subject for psychological research; but an investigation of the closely allied subject of stenography has never been carried on, although the learning of an artificial graphic language is equally well suited to study by similar methods. It was therefore the object of this investigation [1] to determine the form of the learning curve in shorthand by analyzing the records of a great number of students, working under school conditions and motivated by school and business requirements. These students have been studied over a comparatively long time. When such persons are studied for only a short period of time, the observations present only a cross-section of the learning process, and they cannot be said, therefore, to be representative of the whole. Our measurements are restricted to those of a quantitative nature and the tests used were shown to be of a high degree of reliability.

This investigation was started in January, 1926, and ended in April, 1927. The study was made possible by the generous coöperation of the following business schools of New York City: The Katherine Gibbs Secretarial School for the school years 1925–26 and 1926–27; The Miller Business School for the year 1925–26; The Packard Commercial School for the year 1925–26; Miss Conklin's Secretarial School for the year 1926–27; and the U. S. Secretarial School for the year 1926–27. These schools furnished students of four shorthand systems, the Gregg, Burnz, Munson and Pitman. The results for the two last systems are too fragmentary to be included in this presentation, as there were only 26 students of the Munson and about 30 of the Pitman System. The Burnz, Munson, and Pitman Systems are very much alike in fundamental principles, whereas the theoretical basis of the Gregg System is quite distinct; in fact so distinct that one might speak of two great *types* of stenography rather than of various *systems* of shorthand. The Burnz and Gregg Systems are excellent representatives of these two great types of stenography.

[1] The material used was furnished by the Institute of Educational Research of Teachers College, being part of that used in an investigation of adult learning, made possible by a grant from the Carnegie Corporation.

The students of these schools are representative of the larger class of students who attend business schools, and present a wide range in intellect and educational experience. In educational experience the group ranges from grammar school graduates to university graduate students. The median is 12 educational years, or the completion of high school. In intelligence as measured by the Army Alpha the range is from 51 to 194 points with a median at 139.65. In age range, 14–39 years, the group is more narrowly restricted. The median is 19, and the mode is 18 years. Only American-born subjects or foreign-born students who had lived in the United States for a long time are included in the investigation.

The tests used were the *Shorthand Vocabulary Tests*, standardized by E. R. Hoke[2] and published in ten forms, each form being approximately equal in difficulty. Each form consists of 150 of the most common words and phrases in general correspondence and conversation.

In order to avoid the possibility of coaching, these ten forms were rearranged for the present investigation, care being taken to keep the same order of difficulty. A new test, Form A, was made by taking the first, eleventh, etc., words of the original Test 1, of the Hoke series, the second, twelfth, etc., words of Test 2, etc. A new test, Form B, was made by taking the first, eleventh, etc., words of Test 2 (Hoke Series), the second, twelfth, etc., words of Test 3, etc. A new form was given to the student after he had had (approximately) 30 classroom lecture or dictation periods. Through the excellent coöperation of these schools the tests were given under normal conditions of the school room.

All tests were scored by one person. The papers of the respective systems were then checked by persons familiar with the system. The final score was an average of these scores. The basis of scoring was the number of correct items.

The coefficient of reliability of the test was obtained by giving two forms of the test on successive days to students who had had considerable practice. For the Gregg System the coefficient of reliability was $.97 \pm .003$ (P.E.); for the Burnz System, $.92 \pm .01$ (P.E.). It may be contended that this shorthand test is equally reliable as a measure of the knowledge of the Gregg and Burnz Systems. The small difference in the correlation may be explained by the facts that (1) the number of scores (134) on which the reliability coefficient of the Gregg is based is almost twice that of the Burnz scores (73); and (2) the Gregg scores were made by students who had all had approximately the same number of practice hours and who had about reached their limit; this was not true of the students of the Burnz System. These latter students had had widely varying amounts of experience with the subject; some had had less than 100 and others more than 600 practice hours. Hence the Burnz scores are somewhat more 'erratic' in their deviations. (This influence of the amount of practice as a stabilizing factor is shown in Tables I and V.)

The amount of practice in shorthand ranged from 4 hours to more than 600. These estimates included (1) the number of hours spent on shorthand during the school day, (2) the number of hours spent at home in the study of the subject, and

[2] E. R. Hoke, Measurement of achievement in shorthand, *Johns Hopkins Univ. Publ.*, no. 6, 1922.

(3) the number of hours spent on stenography at some other school before entering the school attended at the time of the test. The school attendance record served as a check on (1) and (3); and accurate estimates of (2) were kept by the

TABLE I

GREGG SYSTEM

Time Interval (hours)	Range of Scores Low	High	Mid-Score	Number of Cases
0–9	0	25	5	11
10–19	5	52	17	57
20–	8	67	44	95
30–	14	76	49	12
40–	12	87	48	16
50–	51	113	87	19
60–	23	129	90	19
70–	32	107	92	13
80–	24	137	87	14
90–	43	131	110	43
100–	55	114	90	20
110–	47	130	93	38
120–	47	123	94	35
130–	37	127	101	48
140–	55	136	107	80
150–	65	135	108	51
160–	37	118	94	18
170–	63	136	99	16
180–	35	132	111	17
190–	67	134	120	18
200–	87	132	117	20
210–	89	136	123	21
220–	77	136	113	24
230–	68	136	117	18
240–	69	132	113	18
250–	75	132	119	19
260–	83	132	117	20
270–	62	134	119	22
280–	86	134	116	22
290–	87	133	117	16
300–	98	130	120	9
310–	92	132	117	9
320–	106	132	126	7
330–	109	134	117	5
340–	91	124	107	6
350–	116	124	123	5
360–	106	132	123	4
370–	115	126	123	4
380–	114	135	124	2
390–	111	131	120	4
400–	116	129	121	5
410–	107	134	123	2
420–	108	129	124	4
430–	102	125	114	3
440–	129		129	1
450–	121		121	1
460–	—	—	—	0
470–	126	132	129	3
480–	—	—	—	0
490–	120		120	1

students. The students of one school who constitute the large majority of the total number of students tested did all their studying and practicing in school hours. Students of other schools did varying amounts of study and practice at home.

Learning Curve for the Gregg System of Shorthand

In treating the data the approximate mid-score within practice periods of ten hours was found. These mid-scores, shown in Table I, gave the *general form* of the curve which best represents the data.[3] Table II shows the range of

Table II

Range of Scores from Like Instruction and Like Practice (22 hrs.)

Score	Frequency	Score	Frequency
8	1	47	3
17	1	48	1
20	1	50	2
21	2	51	1
22	1	52	4
29	1	53	1
30	1	54	4
33	3	55	5
34	3	56	2
36	1	57	1
37	1	58	2
38	1	59	2
39	1	60	1
40	2	61	1
41	2	63	1
42	1	64	1
44	1	65	1
45	2	66	1
46	3	67	1

individual differences for a group of students who had all had 22 hours of practice and the same kind and amount of instruction. This table bears out Chapman's [4] contention that individual differences in improvement are just as great as in 'static efficiencies,' and shows the necessity of establishing the

[3] The method of taking the mid-score was used rather than the method of averages because the average weights the extremes too heavily. An inspection of Table I will show that the range of the deviations from the mid-score is very great throughout the lower sections of the curve. The method of averaging would have produced a distorted curve in this section.

[4] J. C. Chapman, The learning curve in typewriting, *J. Appl. Psychol.*, 1919, 3, 252–268.

form of a learning curve by an analysis of records from a great number of students rather than from a few.

A free-hand curve was drawn, using the mid-scores as a general guide, so that approximately 50 per cent. of the experimental points fell on each side of the curve.[5] The values for x and y read from this curve are given in Columns 1 and 2 of Table III. This free-hand curve was smoothed by the method of differencing, and the 'smoothed' values of y are given in Column 5, Table III. These values of y are the ones used in the following computations.

TABLE III

(1) x	(2) 'Free-hand' Value of y	(3) Δy	(4) Smoothed Δy	(5) Smoothed Value of y	(6) $\frac{x}{y}$	(7) Calculated Value of y
0	0			0		0
10	18	18	17	17	.588	18
20	34	16	14	31	.645	
30	46	12	12	45	.667	44
40	55	9	10	53	.750	
50	64	9	9	62	.806	62.5
60	71	8	8	70	.857	
70	78	7	7	77	.909	
80	84	6	6	83	.964	82
90	88	4	5	88	1.022	
100	92	4	4.5	92.5	1.081	91
110	96	4	4	96.5	1.139	
20	100	4	3.5	100	1.200	98
30	103	3	3	103	1.262	
40	105	2	2.5	105.5	1.327	104.5
50	107	2	2	107.5	1.395	
60	109	2	1.5	109	1.467	
70	110	1	1	110	1.545	111.8
80	111			111	1.622	
200	113			113	1.770	117.6
300	119			119	2.521	130
400	122			122	3.278	138

According to Lipka,[6] if a set of data can be approximately represented by an equation of the form

$$y = \frac{x}{a + bx}, \qquad (1)$$

[5] Both the free-hand and smoothed curves split the total number of experimental points (915) into halves. Twenty-one points fall on the smoothed curve (Fig. 1); deducting these from the total, 915, and dividing the remainder by 2, we have 447 points to fall on either side of the curve. By actual count, 441 points fall below the curve.

[6] J. Lipka, *Graphical and mechanical computation*, 1916.

the plot of $\left(x, \frac{x}{y}\right)$ approximates a straight line. This is shown for the values of $\frac{x}{y}$, (Column 6, Table III, and Fig. 1). The method of averaging was used in determining the value of the constants in this equation. The smoothed values of y corresponding to the values of x = 10, 30, 50, 80, 100, 120, 140, and 170 were taken and divided into two groups of 4 each. By averaging the values of the constants, we obtained two simultaneous equations, the solution of which gave the values of

$$a = .501,$$
$$b = .00598,$$

and the corresponding equation

$$y = \frac{x}{.5 + .006x}. \qquad (2)$$

The calculated values of y are given in the last column of Table III.

The curve for this equation is drawn on Fig. 1. This curve also splits the total number of experimental points into halves, approximately 50 per cent. above and below the curve.[7] Hence the curve may be considered an average or 'norm' of improvement in learning the Gregg System of Shorthand.

It is noted from a comparison of Columns 5 and 7, Table III, that the calculated values of y are in very close agreement with the obtained values up to 300 practice hours. After this point, a curve representing the obtained values of y would become rapidly parallel to the x-axis at the value, y = 122; whereas the curve drawn from the calculated values of y shows a diminished but still marked rise after 300 practice hours. It appears from Table I that the number of measures beyond this point is too small to determine whether or not the form of the obtained curve is significant. Many of the

[7] Seventeen scores fall on the curve; deducting these from the total, 915, and dividing by 2, we have 449 scores to fall on either side of the curve. By actual count, 525 scores fall below the curve. Most of this excess is found in the deviation of the calculated curve from the obtained curve after 300 practice hours.

abler students (whose scores helped to determine the form of the curve up to this point) left school before the June closing.

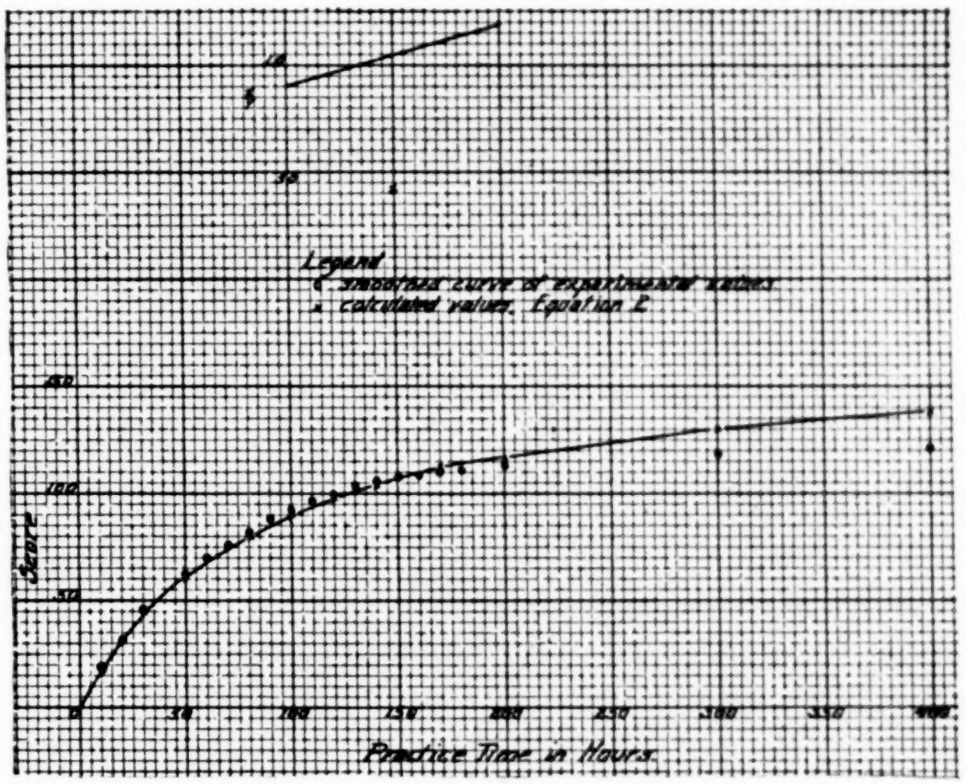


FIG. 1

No simple equation of the form of (1) will give a curve becoming rapidly parallel to the x-axis for some value of y, as can easily be seen from an analysis of the first derivative of equation (1) which is

$$\frac{dy}{dx} = \frac{a}{(a + bx)^2} = \frac{a}{\left(\frac{x}{y}\right)^2}.$$

When the curve becomes parallel to the x-axis, the first derivative = 0, but this is possible only when the ratio, $\frac{x}{y}$, approaches infinity. This limit of the ratio, $\frac{x}{y}$, is, of course, far beyond the range of our experimental data. The calculated curve rises less rapidly after 300 hours, however, and for all practical purposes the influence of practice can be considered negligible beyond this point.

A further confirmation of the applicability to the data of an equation of the form of equation (1) is given by considering the rate of change of slope of the curve. The second derivative of the equation (1) is negative and hence $\frac{dy}{dx}$ decreases as x increases and the curve is concave downward. This is certainly true for the general trend of the results. That it also expresses the trend of the scores of individual students is shown by Table IV, which summarizes four ratios for each of 102 students.

TABLE IV

	Decrease in slope	Increase in slope	Constant
Test 2–1	98.1%	1.9%	None
Test 3–2	63.7%	31.4%	4.9%
Test 4–3	96.0%	2.0%	2.0%

Each ratio was obtained by dividing the difference between the score on one test and the score on the preceding test by the number of practice hours between the two tests. For the first ratio this is obviously the score on the first test divided by the practice time previous to the first test. Only students whose practice had been fairly steady and who had taken four tests were included in this group. Table IV shows the percentage of the total number of students in this group for which the ratios had increased or decreased between the first and second, the second and third, and the third and fourth tests. This table demonstrates that on the average the results show a steady decrease in gain and that a curve which shows negative acceleration throughout is a good representation of the average.

LEARNING CURVE FOR THE BURNZ SYSTEM OF SHORTHAND

The treatment used with the Gregg Shorthand test results was also applied to the Burnz Shorthand scores. The mid-score within each time-interval was determined and served as a guide to the general form of the learning curve. These mid-score values are given in Table V.

A free-hand curve was drawn through these mid-scores so that approximately 50 per cent. of the experimental points fell

TABLE V

BURNZ SYSTEM

Time Interval (hours)	Scores Low	High	Mid-Score	Number of Cases
0–9	—	—	—	0
10–19	—	—	—	0
20–	3	7	5	6
30–	3	8	6	6
40–	3	9	5	21
50–	—	—	—	0
60–	—	—	—	0
70–	7	36	20	4
80–	7	30	15	8
90–	19	46	27	4
100–	7	46	22	10
110–	34	46	40	3
20–	51	54	52	2
30–	25	61	48	3
40–	32	77	53	8
50–	37	49	43	3
60–	47	72	55	5
70–	51	88	71	4
80–	32	68	47	10
90–	48	67	63	7
200–	60	91	75	7
210–	36	97	80	5
20–	44	76	60	4
30–	43	81	55	5
40–	38	98	71	6
50–	41	63	52	3
60–	49	82	67	2
70–	84	97	90	3
80–	65	125	97	7
90–	82	127	90	4
300–	98	112	104	2
310–	66	122	90	8
20–	74	122	96	10
30–	57	126	85	4
40–	88	122	112	9
50–	57	125	92	7
60–	78	126	110	7
70–	108	109	109	2
80–	114	116	115	2
90–	101	131	117	3
400–	87	110	100	4
410–	107	138	117	5
20–	122	124	123	2
30–	88	125	105	5
40–	97	131	123	5
50–	111	132	121	3
60–	112		112	1
70–	114	134	123	6
80–	110	127	120	9
90–	128		128	1
500–	103	125	116	3
510–	132		132	1
20–	123	130	128	4
30–	112	119	116	3
40–	114	129	115	8
50–	121	125	123	2
60–	117	127	122	3
70–	137	141	140	2
80–	126		126	1
90–	130		130	1

above and below the curve.[8] The curve was smoothed by the method of differencing. The 'smoothed' values of y are given in Column 2, Table VI, and the curve obtained is drawn

TABLE VI

(1) x	(2) 'Smoothed' y	(3) $\frac{y}{25.2}$	(4) $\frac{b}{x}$	(5) b	(6) $\frac{b}{x}$ $b = 637$	(7) $\frac{y}{25.2}$ when $b = 637$	(8) y (Cal.) when $b = 637$	(9) y (Cal.) when $b = 600$
0	0						0	
20	2	.079	8.30	166	31.80	< .00005	.001	
40	4.5	.179	7.40	296	15.90	< .001	.025	
60	8	.318	6.70	402	10.60	.010	.25	.68
80	12	.476	6.30	504	7.96	.12	3.0	4.70
100	18	.714	5.55	555	6.37	.40	10.2	13.70
120	25.5	1.010	5.00	600	5.31	.83	20.9	
140	34.5	1.370	4.50	630	4.56	1.32	33.3	39.00
160	46	1.830	3.95	632	3.98	1.83	46.1	
180	59	2.340	3.52	633	3.54	2.30	58.0	
200	71.5	2.840	3.10	620	3.18	2.74	69.0	74.4
220	81.0	3.220	2.80	616	2.90	3.06	77.0	
240	88.5	3.510	2.57	617	2.66	3.37	85.0	
260	94.5	3.750	2.42	629	2.45	3.69	93.0	
280	99.5	3.950	2.28	638	2.27	3.95	99.5	
300	103.0	4.085	2.16	648	2.12	4.13	104.0	108.5
320	106.5	4.230	2.06	660	1.99	4.33	109.0	
340	109.5	4.350	1.97	670	1.87	4.48	113.2	
360	111.0	4.410	1.92	691	1.77	4.64	117.0	
380	113.0	4.480	1.87	710	1.68	4.75	119.8	
400	115.0	4.560	1.82	728	1.59	4.85	122.3	124.8
450	120.0	4.760	1.68	756	1.42	5.05	127.3	
500	122.0	4.840	1.60	800	1.28	5.20	131.0	133.0
∞	150.0	5.96						

in Fig. 2. It is evident from inspection that this curve[9] is quite different from that of the general hyperbolic form which investigators have more or less assumed to be the typical form of the learning curve. It was noticed that the general shape of this curve was similar to that representing the variation of the atomic heat of a solid with temperature. One of the simplest equations for this type of curve was given by Einstein (*Ann. d. Physik*, 1907, 4, 22). This equation

[8] Of the total number of experimental points, 273, 23 fell on this 'smoothed' curve (see Fig. 2). Deducting these 23 scores from the total and dividing the remainder by 2, we have 125 scores to fall on either side of the curve. By actual count, 129 scores fall below the curve.

[9] The possibility of such a learning curve for typewriting was suggested by J. C. Chapman in an article on positive acceleration in learning, published in this JOURNAL, 1916, 1, 494–507.

has the general form

$$y = a \frac{\left(\frac{b}{x}\right)^2 e^{b/x}}{(e^{b/x} - 1)^2}. \qquad (3)$$

Several things are to be noted about this function: (1) as the value of x increases and $\frac{b}{x}$ approaches 0, the function approaches a limiting value equal to a; (2) as x approaches 0 the lower limit of y is 0.

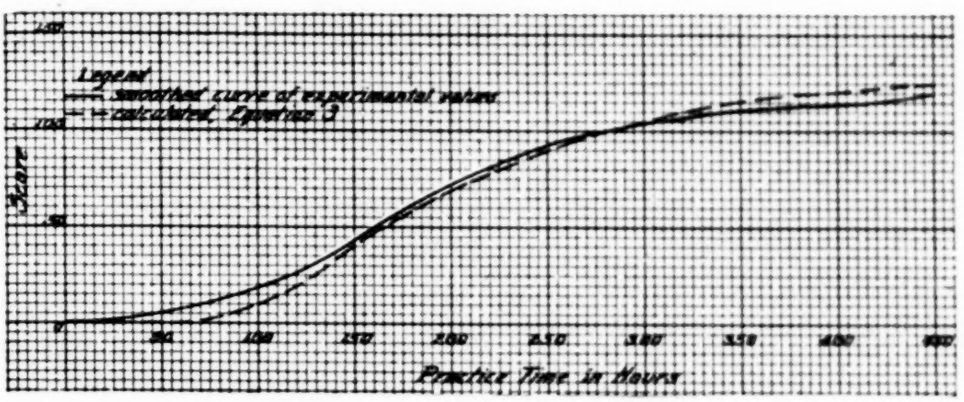


FIG. 2

It is fortunate that this function has been tabulated in several places. The table used here is given in *The new heat theorem* by W. Nernst (Trans. by Barr, 1926.) The table used refers solely to atomic heat. The most convenient way of handling our data was to convert the limiting value to that given in the atomic heat table, namely, 5.96. We have chosen as a limit the highest score which could possibly be made on the shorthand test, namely, 150. The conversion factor is obtained by dividing 150 by 5.96; it is 25.2. All the 'smoothed' values of y were divided by this factor. It was then possible to read the values of b/x directly from the table, or as was actually done from a curve in which b/x was plotted as a function of C_v from the table.[10]

From the known value of x (Column 1, Table VI) and the determined value of $\frac{b}{x}$ (Column 4, Table VI) the value of the constant b was determined. The equation obtained is

$$y = 150 \frac{\left(\frac{637}{x}\right)^2 e^{637/x}}{(e^{637/x} - 1)^2}. \qquad (4)$$

[10] In the atomic heat table the variable T corresponds to our variable x, while the term βv corresponds to our b. It may be noted that βv is a constant for a given solid. The value of C_v in the table will obviously be our value, $\frac{y}{25.2}$.

An average value of b was taken, excluding the two lowest values. It will be seen from Column 6, Table VI, that the value of b itself increases with the values of x, except in the section 120–300 hours, where it is fairly constant. It is in this section of the curve that the most rapid rise is accomplished and which is the most important portion of the curve to be fitted by the equation. It will be seen from a comparison of the smoothed and calculated y values that essentially good agreement is obtained over this range. The change in the value of b through the lower ranges, and in the upper ranges from 300 hours upward, is due more to the way in which the original curve was drawn for y as a function of x than to any other single factor. Had the experimental results for the beginning and final portions of the curve been entirely reliable we should be compelled to consider this variation in the value of b as significant and to abandon the simple Einstein equation. But the time estimates, *i.e.*, x, for the beginning portions of the curve, were known to be of doubtful accuracy;[11] and, at the other end of the curve, the number of results is too small to warrant any great stress upon the significance of the obtained values. If we made the assumption that the experimental values were correct for these ranges, and that the Einstein equation represented the learning curve for the Burnz System, we should have been forced to set a limiting value of approximately $y = 140$ instead of 150, the possible perfect score on the test. Since the limiting value of 150 is by far the more reliable score, it seems more reasonable to use this rather than set a lower limiting value which would intimate that no student of this system could possibly make a perfect score on the test. Through the middle portion of the curve, *i.e.*, from about 140 to 340 practice hours, the values calculated by means of the Einstein equation agree very well with those obtained from

[11] This inaccuracy was due to the fact that all these records were made by students who did varying amounts of study outside of school hours. When the first test was given a number of these students could only estimate roughly the amount of previous work done. After the testing program was initiated more accurate estimates were kept by the students. Also, the data for the range, 0–70 practice hours, were meager, a fact which further affects the reliability of the curve in this section.

the smoothed curve. It will be observed from Table V that for the ranges between 100 and 140, and 220 and 260 practice hours, the amount of data is rather meager and does not warrant any great stress upon the significance of the obtained values for these two periods of 40 hours each.

The curve representing the values of y, calculated by means of the Einstein equation, is shown in Fig. 2. This curve also splits the total number of experimental points into halves.[12]

Although the Einstein equation undoubtedly gives values which are too low for the first 120 hours, it is a better representative of the data than the hyperbolic form. It is to be noted that while the calculated values are low, the highest score made during the first 50 hours was only 9. If the value of b is lowered, we have the effect of pushing the entire curve upward and to the right, though the middle portion of the curve may appear practically unchanged for small changes in the value of b. For example, if b is taken as 600, the lowest value which seems possible from the experimental results, we have the effect of raising both ends of the curve, although the middle portion of the curve represents the experimental values about as well as that obtained when b is taken at the average value of 637. (See the last column of Table VI.) Whether b be taken at 600 or 637, the calculated score at 60 hours will be less than one.

If a curve of the hyperbolic form had been drawn through the greatest number of mid-scores, it would have been necessary to make it intersect the time axis at about 70 hours with the attendant assumption that the student learned absolutely nothing which could be measured by the test during practically the first six weeks of his study. While he may learn very little, it would be absurd to suppose that he learns nothing. What he is learning are the fundamental principles of the subject. His ability to apply these principles may be very low at the beginning of his study, but after about 70 hours he begins to make better and better scores. His rapid rate after 70 hours would have been impossible if he had learned nothing before that point. Although these scores are not as reliable as those that fall in the middle range of the curve, due to the doubtful accuracy of the time estimates, as pointed out above, and due, secondly, to the general unreliability of the extreme scores on any test, however high the reliability coefficient (which for this test was extremely high), it is impossible and incorrect to ignore these scores as would be done if the equation for the hyperbolic curve were assumed to be the better representative of the data.

It is quite probable that by using a somewhat more complex equation, that is, one containing three or four constants, a better agreement could be secured for the range 0–120 hours than that given by the two-constant equation used; but neither the quantity nor the quality of the data available justifies going to these lengths at the present time.

When the acceleration of the curve is considered, it is quite certain that by no stretch of the imagination can one be justified in representing the results obtained on the Burnz Shorthand Tests by the hyperbolic curve which was applicable to the Gregg

[12] Of the total number of experimental points, 8 fell on the curve; deducting these from the total, 273, and dividing the remainder by 2, we have 132.5 scores to fall on either side of the curve. By actual count, 131 scores fall below the curve.

Shorthand Test results. This is readily seen by an inspection of Table VII which shows the change of rate of slope for the Burnz data.

As in the case of the Gregg results, four ratios were found for each one in a group of students whose practice had been fairly steady and who had had four tests. Thirty-eight students or 14 per cent. of the total number were included in this group. Each ratio was obtained by dividing the difference between the score on one test and the score on the preceding test by the number of practice hours between the two tests. For the first ratio, this was the score on the first test divided by the amount of practice time previous to the first test. Table VII shows the percentage of the total number of students in the group for which the ratios had increased or decreased between the first and second, the second and third, and the third and fourth tests.

TABLE VII

	Increase in slope	Decrease in slope	Constant
Test 2–1	63.2%	28.9%	7.9%
Test 3–2	34.2%	50.0%	15.8%
Test 4–3	2.6%	92.1%	5.3%

According to the Einstein equation the acceleration of learning, $\frac{d^2y}{dx^2}$, would be positive in the beginning, passing through a point of inflection, and becoming negative at the end. This is illustrated in Table VII. These results could not be expressed by a curve which was concave downward throughout.[13]

[13] The absence of plateaus in these curves may seem perhaps strange to one familiar with this phenomenon of learning curves. Figs. 1 and 2 represent the average or normal progress made in the learning of shorthand. Such curves could not represent the individual plateaus of every subject since it can hardly be supposed that any two people reach the same plateau at the same time and leave it at the same time. This fact is demonstrated by the overlapping in the range of practice hours required for various stages of learning shorthand. (This point is discussed at greater length in Part III of the Author's thesis, *A study of adult learning of shorthand and typewriting.*) Such plateaus are supposed to indicate that the individual has reached another (higher) stage in his learning process. The number of practice hours required for these different stages varies for each individual. For the Gregg Shorthand, the range for each stage was found to be as follows:

Summary

1. Scores from 915 Gregg Shorthand Tests were plotted on a single graph. The data could best be represented by a hyperbolic curve. The equation for the curve was

$$y = \frac{x}{.5 + .006x}.$$

$\frac{d^2y}{dx^2}$ was negative, showing that $\frac{dy}{dx}$ was decreasing, and the curve was concave downward throughout. For all practical purposes the influence of practice can be considered negligible after 300 hours of study.

2. Scores from 273 Burnz Shorthand Tests were plotted. The equation for the average curve of these data was

$$y = 150\frac{\left(\frac{637}{x}\right)^2 e^{637/x}}{(e^{637/x} - 1)^2}.$$

The acceleration of learning, $\frac{d^2y}{dx^2}$, for this curve is positive in the beginning, passes through a point of inflection and becomes negative at the end. The influence of practice may be considered negligible after 400 hours of study.[14]

Stage A	4—120 hours
Stage B	15—200 hours
Stage C	40—300 hours
Final	80—600 hours

For the Burnz System, the range was as follows:

Stage A	18—225 hours
Stage B	40—500 hours
Stage C	60—600 hours
Final	250—600 hours

There is a marked overlapping in amount of practice required to reach these 'critical stages' or plateaus, when the learning process is studied for students *en masse*. Hence, plateaus in an 'average' curve would be meaningless, since it cannot be shown that every individual requires the same amount of time to attain and complete a higher stage in his learning process.

[14] *Practical Application of These Learning Curves.*—These average curves or 'norms' for the learning of the Gregg and Burnz Systems of Shorthand are of practical value to the teachers and students of stenography. By giving these standardized

tests and by comparing an individual's score with the score made by the 'average' student with the same amount of experience with the subject, the instructor is enabled to make a quantitative and objective estimate of the student's ability and progress. He is further enabled to determine the kind and amount of instruction necessary. When it is recognized that the student has reached his limit, his daily amount of practice can be lessened and he can be given extra instruction in other subjects. The instructor can also make some estimate of the amount of time required to reach a certain degree of proficiency in stenography. These problems are of vital importance to the commercial schools, which are faced both with the practical demands of their students and with the standards and requirements of the future employers of their graduates.

(Manuscript received Oct. 31, 1927)

ORAL PERCEPTIONS IN RELATION TO ANOSMIA

BY H. R. CROSLAND, RUTH C. MILLER AND WINIFRED E. BRADWAY

University of Oregon

In 1926 the senior author published a brief paper [1] upon a case of anosmia which had resulted from a blow on the head incident to a fall approximately one year before his smell sensitivity was tested. Although the tests to which he was subjected were not repeated over and over again from day to day, it seemed evident that his anosmia was complete; that his taste sensitivity had not been affected by the loss of smell; that he still possessed memories for certain smells; and that he was normal or above normal in his ability, by complicated oral perception, to name correctly the food particles and other stimuli which were placed in his mouth while his eyes were shut.

It has been our good fortune to meet another case of anosmia, in a man twenty-two years old, which had resulted from his having been struck on the nose some six years ago. As a result of the accident his nose is somewhat misshapen and his left nostril has been partially obstructed by a piece of obtruding bone. He has reported that he lacks memory of smell for olfactory stimuli antedating the accident. He has steadfastly believed that his taste sensitivity is impaired and has furthermore believed that to judge correctly the nature of food he must rely upon sight. In carrying out experiments upon this subject we have tested his smell, taste and oral sensitivity over and over again with the same strengths of stimulus and, in case of taste, we have gradually increased and reduced the strengths until we have approached the stimulus threshold. Our first object in the experiment, as in the previous one, has been to discover how complete is the anosmia of the subject; the second object has been to work

[1] H. R. Crosland, M. Goodman, & A. Hockett, Anosmia and its effects upon taste perceptions, this JOURNAL 1926, 9, 398–408.

out his threshold of smell in case the anosmia is found to be incomplete; thirdly, we have attempted to determine his thresholds for the four primary tastes, salt, sweet, bitter and sour, to discover whether or not taste sensitivity has been heightened or reduced by the partial or total loss of smell and by the accident itself; and, fourthly, we have sought to discover whether our subject is as adept as the blindfold normal subject in correctly naming what food particles are placed on his tongue, by the use of taste, kinesthesis, pressure and temperature.

The two experimenters, Bradway and Miller (Subjects *B* and *C*), prepared the stimuli and presented them to the subject. They themselves served as controls, each working in ignorance of what stimuli the other was employing. The anosmic subject was blindfold from the time he entered until he left the laboratory room, so that he would in no instance learn of the preparation of the substances to be smelled or tasted and would possess no information upon the labels placed on the bottle containers. The conversation of the two experimenters was limited to an irreducible minimum so that the subject would receive no hints or cues before, during or after the experimental session.

The solutions employed as smell stimuli were placed in small vials of uniform cross-section, with necks approximately such that 77 sq mm of surface was exposed to the nostril. The opening of the vial was inserted into the right nostril, the less injured one, each time for a period of four seconds; it was then withdrawn and the subject reported on the qualities which he had just experienced. During the act of inhalation of the substance the subject held the left nostril tightly shut with one of his fingers. Both the sniffing method and the method of deep inhalation were adopted. The temperature of the room and of the solutions was kept approximately at 20° Cent. during the experiment. The substances were made up in various strengths and were placed in bottles which had been carefully boiled.

TABLE I

THE SMELL STIMULI IN VARIOUS STRENGTHS, AND REACTIONS THERETO BY THE ANOSMIC SUBJECT

	1/500 (Mar. 28)	1/500 (Apr. 28)	1/500 (Apr. 29)	1/500 (May 5)	1/20,000	1/40,000
Picric acid	Nothing	Nothing	Nothing	Weak acid	Nothing	Nothing
Ethyl butrate	Nothing	Nothing	Nothing	Nothing	Nothing	Nothing
Iodine	Nothing	Nothing	Nothing	Nothing	Nothing	Nothing
Camphor	Nothing	Sharp	Nothing	Camphor	Nothing	Nothing
Acetic Acid	Nothing	Pin prick	Pricking	Pricking	Pricking	Nothing
Phenol	Nothing	Nothing	Nothing	Nothing	Nothing	Nothing
NH_4OH	Nothing	Nothing	Nothing	Nothing	Nothing	Nothing

Subject *C* smelled all substances correctly at a strength of 1/500 and nearly all of them at 1/20,000.

Subject *B* smelled nearly all of them correctly at 1/500, one half of them correctly at 1/20,000 and one half of them correctly at 1/40,000.

Tables I and II exhibit the results on the subject's olfactory sensibility, his description of his sensory experiences being recorded in columns opposite the stimulus and under the concentrations used.

TABLE II

SMELL STIMULI AT FULL STRENGTH, AND THE ANOSMIC SUBJECT'S REACTIONS

	April 8	April 9	April 10	April 14
Clove oil	Nothing	Nothing	Nothing	Nothing
Iodine	Acid sensation (throat tonsilitis)	Sensation, don't know what	Vegetable smell; iodine?	Sleepy smell
Camphor	Spreading in nose sensation, refreshing	Alcohol? too strong for alcohol	Sweet smell; camphor?	Dull, mild and sweet
NH_4OH	Burning, stinging, probably ammonia	Strong, like ammonia	Sharp smell, HNO_3	Sharp, like fumes from fumigating
Phenol	Nothing	Sweet, sickening sensation, ether	Flat, faintly sweet, peppermint?	Faint itching
Picric acid	Burning sensation	Sharp smell, nicotine?	Burning smell, grease compound	Biting smell
Ether	No smell, smoky flavor, camphor	Oily taste like HNO_3	Camphor; tastes like castor oil	Like cascarets
Ethyl butrate	Ether, arsenic	Like medicine	Choking, suffocating sensation	Brown, stuffy sensation in throat
Moth balls	Nothing	Nothing	Nothing	Nothing
Chloroform	Sweet taste, cinnamon?	Sweet, penetrating smell	Ether	Sweet taste, like vanilla
Exylene	Nothing	Like camphor	Nothing	Nothing
Acetic acid	Piercing sensation	Pricking, like pins	Dull smell, deadening	Clogging feeling, fills up nose
Formalin	Very burning, probably HCl	Sharp and stinging	Piercing, burns like mustard	Sharp, like medicine
Oil rose	Oily sensation	Nothing	Nothing	Nothing
Oil peppermint	Peppermint	Sweet smell, peppermint	Sweet smell, perfume?	Never smelled before, clearing sensation
Burning Sulphur	Piercing odor, H_2S(?)			

The experiments on the taste sensitivity of the subject were conducted in much the same fashion as those on smell. The subject was blindfold. Drops of various strengths of salt, sugar, quinine and acid solutions were placed by a glass rod on his tongue, the drops being fairly uniform in size. After each such presentation, S named the quality of taste which he had just experienced. The results show about the same gustatory sensibility for him and for the normal subjects.

TABLE III

TASTE STIMULI, AND THE ANOSMIC SUBJECT'S TWO REACTIONS

	May 13th		May 14th		May 17th	
	On tongue	In mouth	On tongue	In mouth	On tongue	In mouth
Potato, raw	Paper	Flat, potato	Cold	Don't know	Cold	Don't know
Sauer kraut	No taste	Burning	Bitter	Sour grass	Sour	Don't know
Cube sugar	Sweet	Sugar	Hard	Sugar	No taste, bean?	Sugar
Bread	Nothing	Don't know	Paper	Bread?		
Plug tobacco	Salty, bitter	Cabbage root	Bitter	Tobacco	Nothing	Tobacco
Raisin	Nothing	Raisin	No taste	Raisin	Sweet	Raisin
Tillmk. cheese	Salty	Don't know	Peppermint	Cheese	Sweet	Bitter
Lemon extract	Burns		Hot			
Vanilla extract	Burns		Vanilla			
Glycerine	Sweet		Chocolate			
Candy (pepm't)	Sweet	Very sweet	Sweet	Like extract	Peppermint	Peppermint
Choc. (Hershey's)			Sweet, choc.	Sweet, licorice	Nothing	Licorice?
Carrot, raw			Cold	Radish	Flat, cool	Carrot
Onion			Don't know	Onion	Onion	Orange peel
Apple			Sweet	Apple	Sweet	Apple
Lemon			Heavy	Lemon	Bitter	Lemon
Orange			Cool, sweet	Orange	Cold, sweet	Orange
Banana			Stinging	Banana	Flat, sweet	Banana
Pickle			Bitter	Sour pickle	Cold	Sour pickle
Beeswax			Sweet	Sweet	Sweet	Don't know
Soap			Flat	Soap	Don't know	Soap
Candied orange peel			Sweet	Sugar in it		
Banana peel			Sweet	Apple peel		
Orange peel			Sweet	Orange peel		

The fourth experiment consisted in presenting various food substances to the tongue of the blindfold subject. The food particles, diced in various shapes, etc., were first placed on the tongue while it was extended, left there for two seconds and then removed, whereupon the subject attempted to name the food substance and to describe in detail what he had tasted and felt. The series of food particles was also presented a second time. This time the subject was allowed to withdraw the tongue and food into the oral cavity and to roll the food around on his tongue but not to bite or to swallow it. The food was then removed from his mouth and again he reported upon what he believed the substance to be, and in detail what he had tasted and felt. The results from the fourth experiment are shown in Table III.

Conclusions

1. The subject's anosmia is nearly or quite complete. The only substances which he appeared to smell at all were peppermint and chloroform, the latter not being known as chloroform but reported as 'tasting sweet.' It is not clear that he actually smelled these substances; their 'pungency' and taste were apparently effective. They were reacted to correctly only when in a very strong, undiluted form. The other substances caused kinæsthesis, temperature and pressure from the mucous membranes of the nasal cavity, described as burning or pricking. All of these sensory experiences, which are accessory to smell proper, were lost with dilutions of acetic acid and camphor of the strength of 1/500. The pricking from acetic acid was retained in dilutions as weak as 1/20,000.

2. The conclusion from the above is that the kinesthetic, thermal and pressure qualities are usually lost in our anosmic subject at dilutations of 1/500. The control subjects had not only the kinæsthetic and other experiences but also they accurately smelled the substances at 1/500 and a great many of them at a strength of 1/20,000.

3. The subject's taste sensitivity appears to be neither above nor greatly inferior to that of the two controls in this

experiment and of the two who served in the experiment of 1926. His ability to detect the presence of salt is slightly inferior to that of the controls, while his sensitivity for sweet is slightly superior to theirs; they all possess equal sensitivity for quinine sulphate.

4. In judging the nature of food substances placed on his tongue the anosmic appears to be as capable as the controls, but certainly not superior to them. Without tongue movements and salivation, with their attendant kinæsthetic and pressure stimulations, he could not with accuracy describe the substances. His belief that his oral abilities in perception are below normal is unfounded.

5. The wide daily variation in reports suggests the necessity for many determinations in obtaining thresholds of taste and smell.

(Manuscript received Oct. 8, 1927)